The Dissertation Committee for Sari Andoni Certifies that this is the approved version of the following dissertation:

Motion Selectivity as a Neural Mechanism for Encoding Natural Conspecific Vocalizations

Committee:

George D. Pollak, Supervisor

Jonathan Pillow

Ila Fiete

Nicholas Priebe

Rafael de la Llave

Motion Selectivity as a Neural Mechanism for Encoding Natural Conspecific Vocalizations

by

Sari Andoni, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2010

To my family, friends, and beloved ones.

Acknowledgements

I would like to thank my mentor Dr. George D. Pollak for introducing me to the wonderful world of the auditory system. He has provided me with his generous support, encouragement and guidance throughout my years in his lab and I'm very grateful for gaining immense amounts of scientific knowledge from his many years of great science.

I would also like to thank Drs. Jonathan Pillow, Nicholas Priebe, Ila Fiete and Rafael de la Llave for agreeing to serve on my committee and providing me help and guidance in my research.

My thanks also go to members of my lab, Dr. Josh X. Gittelman for teaching me about biophysics and intracellular recordings, Na Li for performing the iontopheresis and extracellular recordings in my first study and for teaching me how to do electrophysiology, Carl Resler for his technical and social support, as well as Dr. Ruili Xie for helpful conversations and advice.

I would also like to thank other people in the department for their help and support, including Alfredo Ghezzi, Yazan Al Hasan and Jascha Pohl.

My thanks also go to my friends and fellow musicians in the city of Austin for making my graduate life a memorable one.

Motion Selectivity as a Neural Mechanism for Encoding Natural Conspecific Vocalizations

Publication No._____

Sari Andoni, PhD

The University of Texas at Austin, 2010

Supervisor: George D. Pollak

Natural sound, such as conspecific vocalizations and human speech, represents an important part of the sensory signals animals and humans encounter in their daily lives. This dissertation investigates the neural mechanisms involved in creating response selectivity for complex features of natural acoustic signals and demonstrates that selectivity for spectral motion cues provides a neural mechanism to encode communication signals in the auditory midbrain. Spectral motion is defined as the movement of sound energy upward or downward in frequency at a certain velocity, and is believed to provide the auditory system with an important perceptual cue in the processing of human speech. Using the Mexican free-tailed bat, *tadarida brasiliensis*, as a model system, this research examined the role of selectivity for spectral motion cues, such as direction and velocity, in creating response selectivity for specific features of the social communication signals emitted by these animals. We show that auditory neurons

in the midbrain nucleus of the inferior colliculus (IC) are specifically tuned for the frequency-modulated (FM) direction and velocities found in their conspecific vocalizations. This close agreement between neural tuning and features of natural conspecific signals shows that auditory neurons have evolved to specifically encode features of signals that are vital for the survival of the animal. Furthermore, we find that the neural computations resulting in selectivity for spectral motion are analogous to mechanisms observed in selectivity for visual motion, suggesting the evolution of similar neural mechanisms across sensory modalities.

Table of Contents

List of Figures

Chapter 1: Introduction

Most animals, including man, have evolved to communicate with members of their own species. Conspecific communication facilitates collective social behaviors that allow the survival and well being of the species. Whether in the behavior of a male trying to court a female, an infant calling his mother, or an alarm call warning of an incoming predator, social communication is an essential part of the everyday behaviors of most animals. Social communication took on different forms across different species and had led to the evolution of complex systems of motor control as well as sophisticated sensory systems. These systems gave animals the ability to emit detailed communicative signals containing meaningful behavioral relevance. However, these complex signals would be insignificant without a sensory system that turns each communication signal into a neural perception which guided the animal to engage in a specific social behavior. Yet communication in each species was never restricted to a single sensory modality but it rather utilized the sensory and motor abilities that were used in other behaviors. This is evident in the examples of chemical signaling in simple organisms (Waters and Bassler 2005), the flight dances of honeybees (Riley, Greggers et al. 2005), the flapping of wings in the courtship behavior of the fly (Ewing and Bennet-Clark 1968), or the facial expressions of humans (Morris, Frith et al. 1996). However, the most complex and advanced form of social communication is found in animals that produce a social signal encoded in a pattern of air (or water) vibrations termed a vocalization. Communication through sound is present in a variety of species across both vertebrates, such as birds and

most mammals, as well as invertebrates, such as frogs, flies and even mosquitoes (Gibson and Russell 2006). Perhaps the most evolved form of acoustic communication is found in human speech with complex features such as meaning and language.

To investigate how sensory systems could process these complex social communication signals, this dissertation examines the neural mechanisms used by the auditory system to process natural sounds, such as conspecific vocalizations and human speech. Using the auditory midbrain of the bat as a model system, this research explores the mechanisms in which auditory neurons encode specific spectrotemporal features of the acoustic signals these animals use for social communication. We begin by revealing the importance of sounds that move upward or downward in frequency as the main characteristic of the acoustic signals of bats. The significance of this motion across frequency is then related to our own perception of human speech. We then explore mechanisms of motion selectivity observed in other sensory modalities and later relate them to mechanisms observed in the processing of conspecific communication signals by the auditory system.

BAT VOCALIZATIONS

To understand the neural mechanisms involved in acoustic social communication, many researchers have studied the motor control used in vocal production while others concentrated on the neural processing of sound in the auditory system. Even though most early studies of social communication both in vocal production and neural processing had concentrated on birds and their learned courtship song (Doupe and Kuhl 1999; Mindlin and Laje 2005; Theunissen and Shaevitz 2006), recent attention had been given to the

social communication signals of bats (Bohn, Schmidt-French et al. 2008; Bohn, Schmidt-French et al. 2009). For decades, bats have been extensively studied for their acoustic specialization in their echolocation behavior (Suga and Schlegel 1972; Simmons, Fenton et al. 1979; Pollak and Casseday 1989), yet it was only recently discovered that they possess a large repertoire of social communication signals (Bohn, Schmidt-French et al. 2008). Their social vocalizations formed a large library of calls that were emitted when the animal was engaged in a specific social behavior such as feeding, courtship or mother-infant interactions. Figure 1.1 shows a spectrogram for a set of example calls from Mexican free-tailed bats. A spectrogram is constructed by plotting the amplitude of the windowed Fourier transform of overlapping time segments of the sound waveform, shown at the bottom of the spectrogram. Each call was vocalized under the context of an observed social behavior as illustrated in the figure. Examining the features of bat vocalizations, one can observe rapid changes in amplitude over time, termed amplitude modulation (AM), as seen in the courtship trills. More importantly, it is apparent that some of the more complex features of bat communication signals included a prominence of frequency-modulated (FM) sweeps. FM sweeps are acoustic signals that move upward or downward in frequency at a specific velocity.

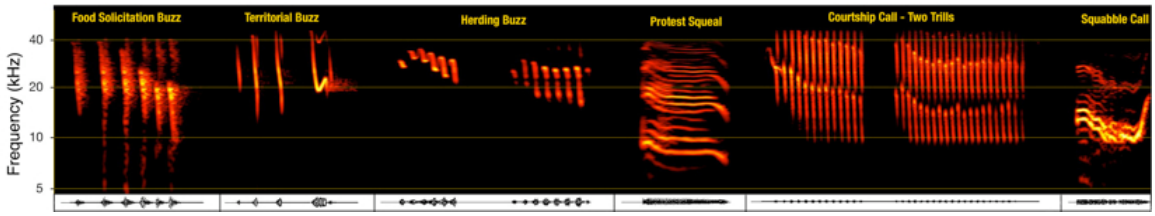

Figure 1.1: Social Communication Signals of Mexican free-tailed Bats.
A spectrogram of a set of six calls recorded from a local colony of Mexican free-tailed bats. The main features of these calls mainly consisted of frequency modulations as observed in the territorial buzz while some had strong amplitude modulations as seen in the courtship trills.

SPECTRAL MOTION

Acoustic signals that sweep upward or downward in frequency at various speeds are not limited to the social communication signals of bats. In fact FM sweeps can be observed in the vocalizations of most, if not all, animals including human speech. This movement of sound energy across the frequency axis is referred to here as spectral motion. Although the concept of motion is typically applied to objects moving in *space* both in visual and auditory modalities, upward and downward FM sweeps are considered here a form of motion for two main reasons. First, when an auditory signal moves across frequency it results in a movement of mechanical energy across the sensory surface of the cochlea in a similar manner to a visual object moving across the sensory surface of the retina. Second, in human listeners, FM sweeping sounds produce the perception of motion across successive pitches both in the rising or falling directions. Therefore, by having both the physical and the perceptual hallmarks of motion, a sound that moves across frequency might present the auditory system with an important sensory cue that could participate in building a perception of a behaviorally relevant signal such as a social communication call.

SPECTRAL MOTION IN HUMAN PERCEPTION

In human speech, spectral motion takes the form of formant transitions. Formants represent the resonances of the vocal tract during speech, and formant transitions occur when consonant and vowel gestures are combined to form syllables (Liberman and Mattingly 1989). Figure 2.1a shows the formants and their transitions as a result of articulating the syllables [da] and [ga]. When the third formant starts with an upward sloped formant transition, the syllable [ga] is perceived, whereas a transition in the downward direction results in the perception of [da]. As can be seen in the figure, formant transitions are similar to upward and downward FM sweeps and are vital for the perception of different vowels and consonants depending on their spectral location, direction and velocity (Lindblom and Studdert-Kennedy 1967; Stevens and Klatt 1974).

To further illustrate the significance of spectral motion in the perception of human speech, psychophysical studies have demonstrated that listeners are able to perceive a given speech sound by listening to a group of tones whose frequencies were modulated to match the formants in the original speech utterance. They termed the tonal stimuli sine-wave speech (Remez, Rubin et al. 1981). Figure 1.2b-c shows a spectrogram of a male speech sample and its sine-wave speech reproduction. By matching the movement of energy across frequency, with a slight modulation in amplitude, listeners are able to perceive the speech utterance in the original sample. This shows that the auditory system must be using spectral motion cues in the processing and perception of human speech.

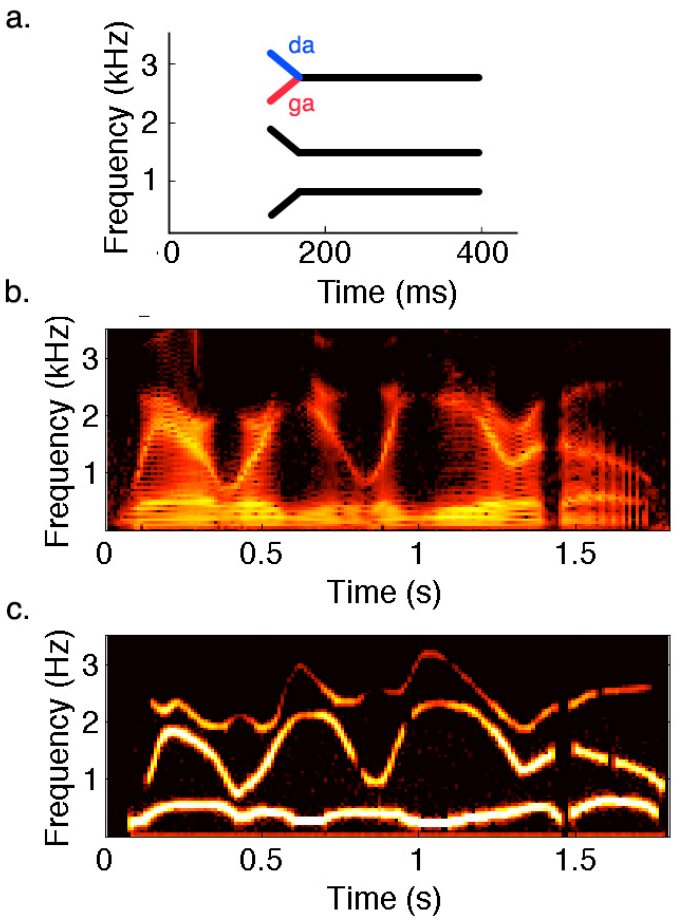

Figure 1.2: Spectral Motion in Human Speech.
(a) The plot shows the three formants present in the syllables [da] and [ga]. An upward transition of the third formant (red) results in the perception of the syllable [ga], whereas a downward formant transition (blue) leads the listener to perceive [da]. (b) A speech sample of a male saying, "Where were you a year ago?" (c) Sine-wave speech is created with three tones moving in frequency to match the formants and their transitions found in b. With the addition of some amplitude modulation, a human listener perceives the same speech sentence the speaker articulated in b.
(a is redrawn from Liberman and Mattingly (1989) and b,c are adapted from Remez, Rubin et al. (1981), both with permissions.)

In addition to speech perception, psychophysical studies have shown that the human auditory system responded to both first and second-order spectral motion, even though second-order motion did not involve any obvious displacement of energy throughout the stimulus presentation but rather had a spectrally moving notch (Huddleston and DeYoe 2003). Furthermore, studies had also revealed that spectral motion produced an auditory after-effect similar to that in the visual "waterfall" illusion; where a prolonged exposure to a moving stimulus resulted in perceiving a stationary stimulus as moving in the opposite direction (Shu, Swindale et al. 1993; Dawe, Platt et al. 1998).

MOTION DETECTION IN VISION

Motion perception is not an exclusive property of human vision and audition. In fact, motion detector neurons are found in a wide variety of animals from invertebrates to mammals, and across most sensory modalities including somatic sensation. Motion detection may have evolved in sensory systems to detect changes in the environment whether in avoiding predators, chasing prey or any change that might affect the overall survival and wellbeing of the animal.

In order to classify a sensory neuron as a motion detector, it has to respond differentially to the direction of a moving stimulus. Hubel and Wiesel (1959, 1962) first found neurons in the primary visual cortex (V1) of the cat that were selective for the direction of visual motion. Reichardt (1961) studied the behavior of the beetle in response to moving stimuli and extracted a model that later explained motion selectivity observed in visual neurons of the fly (Borst and Egelhaaf 1989; Steveninck and Bialek 1995).

Barlow and Levick (1965) examined direction-selectivity in the retina of the rabbit and studied the effect of inhibitory inputs.

While many models have been suggested to explain the neural mechanism that enables a visual neuron to be motion selective, they differed significantly in their underlying computation. The Reichardt model (Reichardt 1961), one of the earliest models of motion detection to be formally described, relied on computing the correlation between two adjacent spatial points with different temporal filtering (delays). It is thus referred to as the correlation-type model (Figure 1.3a). The full model consisted of two mirror-symmetric directional subunits: an excitatory subunit providing facilitation in the preferred direction, and an inhibitory subunit providing suppression in the non-preferred direction (Figure 1.3b). Although Reichardt described each subunit as being nonlinear, correlating its inputs using multiplication (Egelhaaf, Borst et al. 1989), a similar result could be obtained if both subunits were linear integrators with a thresholded output

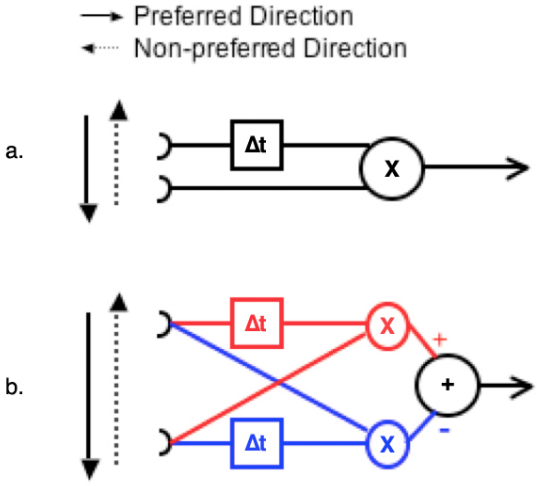

Figure 1.3: The Reichardt Correlation-type Model for Motion Selectivity. (a) A direction-selective subunit is constructed by temporally filtering an input with a certain delay (Δt), followed by correlation through multiplication. (b) The full Reichardt model consists of two opposing directional subunits, one providing facilitation in the preferred direction (red) with the other suppressing the response in the non-preferred direction (blue).

(Grzywacz and Koch 1987). In either case, the direction-selective output is nonlinear in reference to the stimulus.

Another prominent model for motion selectivity was described by Adelson and Bergen (1985) and is known as the energy model. In their model, motion energy is computed using at least two linear filters that are inseparable and tilted in the spatiotemporal plane. The filters are aligned to be in quadrature phase (90°) and their output squared and summed to produce a motion selective output. Under certain conditions, such as contrast normalization, the computation performed by both the Reichardt correlator and the energy models is found to be equivalent (Adelson and Bergen 1985).

A third class of models for motion selectivity depended on the spatial and temporal gradient of the stimulus (Limb and Murphy 1975; Fennema and Thompson 1979; Hildreth and Koch 1987). The gradient model computes the velocity of motion directly by dividing the temporal derivative of the stimulus ($\partial s(x,t)/\partial t$) by the spatial derivative ($\partial s(x,t)/\partial x$). Different from both the correlator and the energy models, the gradient model provides a signal that is proportional to the velocity of motion regardless of the exact spatial pattern of the moving object (Borst 2007). While the gradient model had been popular in machine vision, it did not receive much experimental evidence in biology.

All models for motion detection described above rely on an explicit nonlinearity in computing a motion-selective output. Yet it is still debatable whether a directionally selective neuron integrates its inputs linearly with an output threshold, or rather direction-

selectivity is an emergent property of nonlinear interactions among excitatory and inhibitory inputs. While recent studies of direction-selective simple cells in V1 showed strong evidence for linear summation of inputs both in extracellular (Reid, Soodak et al. 1991) and intracellular recordings (Priebe and Ferster 2005), both recent and early studies of motion selectivity in the retina showed dependence on the nonlinear interaction of synaptic inputs.

In studying ganglion cells in the retina of the rabbit, Barlow and Levick (1965) suggested that inhibition provided a shunt or "vetoed" the effect of excitation when the stimulus moved in the non-preferred direction. In contrast, moving in the preferred direction the shunting inhibition arrived too late to have an effect. Therefore, their model relied on nonlinear suppression in the non-preferred direction with little or no facilitation in the preferred direction. Recent studies also showed that amacrine starburst cells in the retina relied on nonlinear integration of inputs in computing the direction of motion such as dendritic properties and voltage-gated channels (Taylor, He et al. 2000; Euler, Detwiler et al. 2002).

While many more models described the processing of apparent motion in the visual system using other nonlinear mechanisms, less attention had been given to the mechanisms of motion detection in the auditory system.

SPECTRAL MOTION SELECTIVITY IN AUDITION

While an auditory stimulus could have apparent motion both in space and in frequency, each motion axis presents the auditory system with a different processing cue. Spatial motion indicates "where" the sound lies in space as it provides localization

information on the spatial direction and velocity of a moving sound source, whereas spectral motion describes changes in the frequency content of an acoustic signal and may therefore provide important information on how the signal was initially vocalized and "what" its behavioral meaning might be. As described in the previous sections, spectral motion is prominent in most animal vocalizations and plays an important role in the perception of human speech. This has led many to study neural responses to FM sweeps in various parts of the mammalian auditory system.

The first processing station known to have a prominence of selectivity for the direction of spectral motion is the nucleus of the inferior colliculus (IC) in the auditory midbrain. The IC receives a convergence of both excitatory and inhibitory projections from almost all lower nuclei in the brainstem (Figure 1.4). This convergence of inputs results in the emergence of novel response properties that encode various neural computations, which are then relayed to the auditory cortex, via the thalamus, for further processing. In addition to selectivity for spectral motion, emergent properties of IC neurons include duration tuning (Casseday, Ehrlich et al. 1994; Covey, Kauer et al. 1996), combination sensitivity (Mittmann and Wenstrup 1995; Portfors 2004), tuning for multiple binaural cues (Chase and Young 2006) and feature selectivity for natural signals such as conspecific vocalizations (Klug, Bauer et al. 2002; Holmstrom, Roberts et al. 2007).

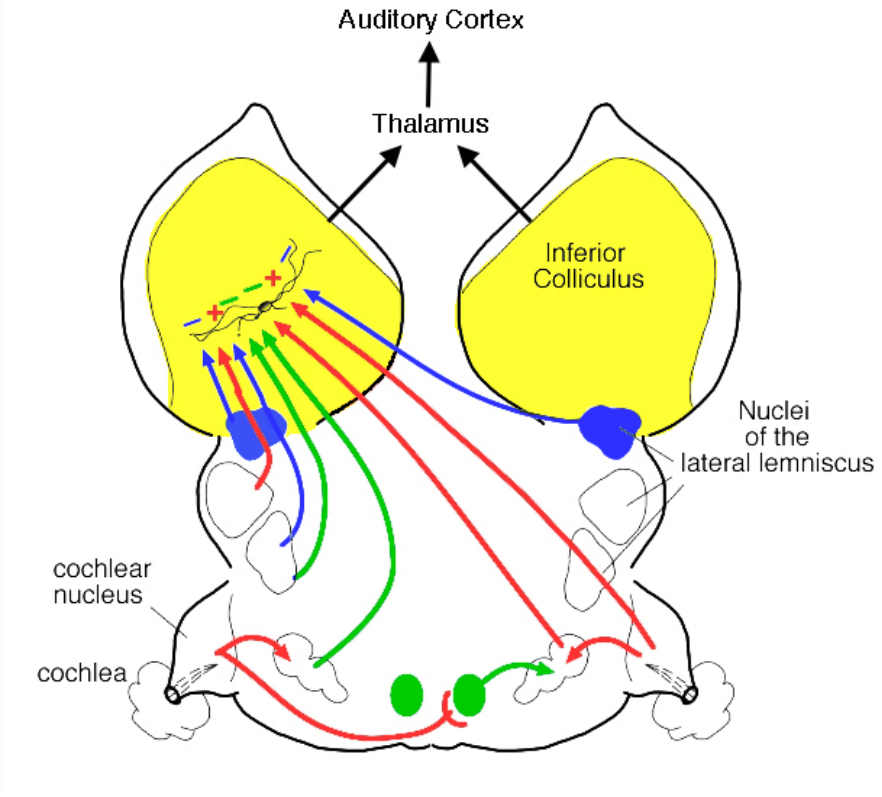

Figure 1.4: Convergence of Excitatory and Inhibitory Projections onto the Inferior Colliculus.
The projections from most lower nuclei in the brainstem converge onto the inferior colliculus (IC). The projections are both excitatory (red) and inhibitory. Some inhibitory projections are glycinergic (green) while others are GABAergic (blue). The IC in turn relays its processed information to the auditory cortex via the thalamus.

Although directional selective neurons were primarily found in the IC as well as the primary auditory cortex (AI), it is still unclear where in the auditory system direction-selectivity is first created. Studies of the eighth auditory nerve showed similar firing rates in response to ascending and descending FM sweeps, yet the discharge patterns differed slightly (Sinex and Geisler 1981). In the cochlear nucleus, some neurons responded differently to the direction of spectral motion, but their selectivity appeared much weaker than seen in the IC or the cortex (Britt and Starr 1976).

Suga (1965) was the first to characterize directionally selective neurons in the auditory system of the bat. He was also the first to suggest that FM direction-selectivity could be a result of surround inhibition (Suga 1968). More recently, direction-selectivity in the IC was shown to be significantly reduced by the blocking of GABAergic (Fuzessery and Hall 1996) and glycinergic inhibitory receptors (Koch and Grothe 1998). In the auditory cortex, *in vivo* whole-cell recordings from AI cells showed that direction-selectivity was already present in their synaptic inputs but was notably enhanced by spectrotemporal asymmetries of cortical inhibition (Zhang, Tan et al. 2003). Although synaptic inhibition seems to play an important role in shaping direction-selectivity of auditory neurons, its exact role and interaction with excitation to produce a directionally selective response is still undetermined.

FEATURE TUNING OF AUDITORY NEURONS

Previous studies have shown that response selectivity for natural communication signals can be observed as early as the inferior colliculus (IC) in the auditory midbrain (Klug, Bauer et al. 2002; Portfors 2004; Holmstrom, Roberts et al. 2007). Receiving a

convergence of excitatory and inhibitory inputs from the brainstem, as was illustrated in Figure 1.4, it is not surprising that emergent response properties like feature selectivity can arise in the IC. It has been previously shown that blocking inhibitory inputs greatly reduced response selectivity to natural signals in the IC (Klug, Bauer et al. 2002), yet it is still unclear which spectral and temporal features of conspecific vocalizations are encoded by an IC neuron and what neural mechanism IC cells are using to produce a feature selective output.

In order to evaluate the relationship between selectivity for spectral motion cues and selectivity for social communication signals, we used the Mexican-free tailed bat as a model system and compared the response properties of auditory neurons in the IC to natural bat calls with their responses to synthetic stimuli having a certain direction and velocity of motion. Our goal was to assess the degree to which we could predict the neural response to natural stimuli based on response features extracted from simpler synthetic stimuli. In contrast, we also evaluated the ability of the functional models we extracted from responses to natural calls in predicting the responses to stimuli synthesized to specifically test the motion selectivity of an IC neuron.

In the first study described in the second chapter, we used a family of moving ripple stimuli to extract the spectrotemporal tuning of each IC neuron. A moving ripple stimulus consists of broadband noise that is sinusoidally modulated both in frequency and in time (Figure 1.5). Similar to visual gratings, each moving ripple contains a certain spectral modulation rate measured in cycles/octave, and a temporal modulation rate specified in Hz. The combination of temporal and spectral modulations gives rise to

upward or downward spectral motion at a velocity equivalent to the ratio of its modulation rates. Since moving ripples make up the Fourier basis components of the spectrotemporal domain, by using a large set of ripples covering a wide range of spectral

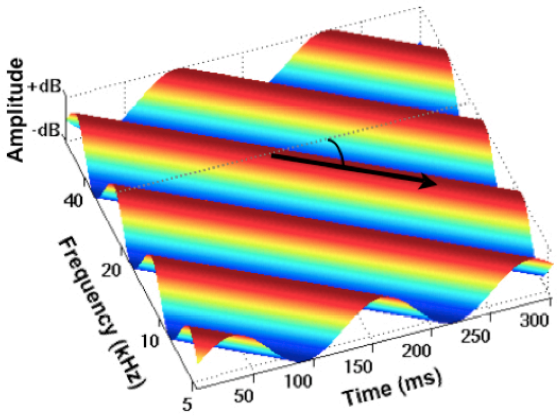

Figure 1.5: Moving Ripple Stimulus. Analogous to visual gratings, a moving ripple consists of broadband noise that is modulated sinusoidally in frequency and in time. The moving ripple shown has a spectral modulation rate of 1 cycle/octave, and temporal modulation rate of 8 Hz, giving rise to downward spectral motion at a velocity of 8 octaves/s (arrow).

and temporal modulation rates, we can derive a linear estimate of the receptive field of each IC neuron using reverse correlation (Klein, Depireux et al. 2000). The spectrotemporal receptive field (STRF) of each neuron showed the neural tuning to specific excitatory as well as suppressive regions in the frequency-time plane. In a minority of IC neurons, the STRF was able to accurately predict neural responses to natural communication calls. Most STRFs with high predictability showed strong spectrotemporal inseparability and were tilted indicating a preference for the direction and velocity of spectral motion. When inhibitory receptors were blocked, the inhibitory regions in the STRF as well as its directional preference were greatly reduced, suggesting a major role for inhibition in shaping directional selectivity in the IC. Comparing the velocities found in the conspecific social signals with velocities IC neurons are tuned for showed close correspondence indicating that IC neurons of the bat are tuned, partially by

inhibition, to respond to spectral motion cues present in their conspecific communication signals. While the STRF was successful in describing the response properties of a minority of IC neurons, the majority of these cells exhibited strong nonlinearities not captured by the linear STRF model or they were not very responsive to broadband stimuli such as moving ripples.

In the third chapter, the spectrotemporal features each neuron is encoding were extracted from responses to bat vocalizations directly, without relying on broadband synthetic stimuli such as moving ripples. By using the linear-nonlinear cascade model (Sharpee, Rust et al. 2004; Simoncelli, Pillow et al. 2004; Bialek and de Ruyter van Steveninck 2005), we were able to derive multiple stimulus features each neuron is selective for, together with the nonlinear relationship of each feature to the response of the IC neuron. This method allowed us to derive the overall receptive field of most IC neurons sampled with higher predictability than using the moving ripple stimulus. Similar to our previous study, the features to which each IC neuron was tuned showed strong selectivity for spectral motion. Comparing the extracted features and their nonlinearities with existing functional models of motion selectivity showed two distinct neural mechanisms for motion selectivity in the IC. Some IC neurons sampled were tuned for cooperating features that could be described accurately by the energy model for motion detection described above, while others were tuned for opposing features in an equivalent arrangement to that described in the Reichardt model. This suggests the evolution of common neural mechanisms across sensory modalities for the detection of apparent

motion and further illustrates the importance of spectral motion cues in the processing of social communication signals.

The last chapter of this dissertation concludes our results and suggests future directions for this research.

Chapter 2: Mapping Receptive Fields with Moving Ripples

In order to study the neural mechanisms involved in the processing of natural communication signals by IC neurons, we must first map their sensory receptive fields. Each IC neuron is tuned for a range of frequencies that might excite or inhibit the response of the cell depending on their intensity and spectrotemporal location. Using a variant of reverse correlation, which derives the receptive field of a neuron by averaging all the stimuli that preceded its spiking response (Marmarelis and Naka 1972), we can map the spectrotemporal receptive field (STRF) of an IC neuron and extract its feature tuning. We can then predict the response of a neuron to natural signals and study the neural mechanisms that enabled each cell to be selective for a particular feature of conspecific communication signals.

To derive the STRFs of IC neurons, we used moving ripple stimuli and evaluated how the spectral and temporal features of inhibition interact with excitation, and how those interactions create selective features for the direction and velocity of FM sweeps, acoustic features that are prominent in their natural calls. To do this we first generated STRFs with a family of moving ripples with a wide range of spectral and temporal modulation rates (Klein, Depireux et al. 2000). We verified the validity of the STRFs by comparing their features both with tone evoked responses and by convolving them with a suite of conspecific communication sounds to evaluate the degree to which the STRF predicts the responses to biologically relevant sounds that were not used to generate it. In a subset of cells we also generated STRFs before and while inhibition was blocked, and evaluated how the spectrotemporal features of excitation and inhibition interact to create selectivity for spectral motion. Finally, we decomposed the conspecifc calls into their ripple components and show that the velocities in the FM components of species-specific

communication calls correspond closely with the FM velocities preferred by IC cells. In short, we show how the spectrotemporal arrangement of inhibition in receptive fields shapes and tunes response selectivity for complex signals and argue that those features are tuned to the natural signals these animals receive in their daily lives.

METHODS

Acoustic Stimuli

Acoustic signals were tone bursts, 6 species-specific calls, and 147 moving ripples. The calls were selected from a larger repertoire and were chosen because their acoustic features represent a range of spectrotemporal patterns that are used in a variety of important behavioral contexts. All stimuli were presented monaurally to the ear contralateral to the IC with custom made earphones (Schuller 1997) biased with 200 V DC and positioned in the funnel formed by the bat's pinna. The earphones were flat $\pm$ 5 dB from about 10 kHz to 70 kHz. The tone bursts, 20 ms in duration with 0.2 ms rise-fall times, were generated by a Macintosh G4 computer with custom-built software. The ripples were created using MATLAB and were stored as sound files. All stimuli were then uploaded from the Macintosh G4 into a custom made Downloadable Arbitrary Waveform Generator through a 24-bit digital interface (National Instruments DIO-24) and a digital distributor just prior to that particular sound's presentation. The acoustic signals output to an Instrutech 16-bit D/A converter and were sent to custom made electronic attenuators and then to the earphones.

The best frequency (BF), or frequency to which it was most sensitive, and threshold at BF were obtained for each cell, followed by a rate-level function and tuning curve. The ripples were then presented at 30-50 dB above BF threshold and a spectro-temporal receptive field was constructed as described below. Peri-stimulus time (PST) histograms were generated from 20 presentations of each stimulus (bin width was 1.0 ms).

The species-specific communication calls were recorded from a captive colony of Mexican free-tailed bats while the bats were engaged in a particular behavior. The calls were equated for RMS intensity and presented to each neuron at 30-40 dB above BF threshold. The calls are shown in Figure 2.3.

Moving Ripple Stimuli

Each moving ripple stimulus was constructed from a combination of tone components that had random phases (Kowalski, Depireux et al. 1996). The tone components were sinusoidally modulated both spectrally and temporally. The amplitude of each tone component was sinusoidally amplitude modulated in time with a fixed rate and a phase that depended on the position of the tone component on the spectral axis. The initial phases of the amplitude modulations of the tone components varied sinusoidally on the spectral axis at a fixed rate thereby creating spectral modulations at any point in time. The particular combination of spectral and temporal modulation rates, in turn, caused frequency modulations, where the velocity, direction and spectral bandwidth of the frequency modulation was unique to each ripple stimulus. Thus, some ripples had relatively narrow frequency bands that swept upward or downward in

frequency at a high rate, whereas others had wider bands that swept at lower rates (Fig.1A-B).

Each tone component of a ripple stimulus can be described as follows:

$$S(t,x) = L \cdot [1 + \Delta A \cdot \sin(2\pi \cdot (\omega t + \Omega x) + \Phi)] \quad (2.1)$$

where x is the position on the logarithmic frequency axis defined by $x = log_2(f / f_0)$, f_0 being the starting frequency, ω is the temporal modulation rate (TMR) in Hz, Ω is the spectral modulation rate (SMR) in cycles/octave, Φ is the starting phase of the ripple, L is the overall loudness of the stimulus and ΔA is the amplitude modulation of the ripple. Each ripple stimulus had 25-100 tones per octave equally spaced along the logarithmic frequency axis and was 90% linearly modulated (ΔA = 0.9). L was typically 30-50 dB above BF threshold.

A total of 147 ripple stimuli were used that were temporally modulated (ω) from 8 to 392 Hz in steps of 64 Hz, and spectrally modulated (Ω) from -3.0 to 3.0 cycles/octave in steps of 0.3 cycles/octave. Each ripple was 300 ms long and had a frequency range of 5 to 50 kHz (~3.3 octaves).

Data Analysis

Each ripple stimulus was presented 10-20 times and a PST histogram generated from the 10-20 repetitions of each ripple was obtained (Figure 2.1C). The PST histogram was then folded into a 16-bin period histogram based on the period of the ripple stimulus (reciprocal of the temporal modulation rate, $1/\omega$). To find the magnitude and phase of the response, the period histogram was fitted with a sinusoid by applying a 16-point Fourier

transform and taking the magnitude and unwrapped phase of the fundamental (Kowalski, Depireux et al. 1996). Response magnitude and phase were extracted for every ripple stimulus presented with different spectral and temporal modulation rates (ω, Ω).

We graph the response magnitude and phase response calculated from the response evoked by each ripple in two matrices, one showing the response magnitudes and the other the response phase (Fig 1D-E). The matrix of the response magnitudes evoked by each ripple is termed the magnitude matrix and the graph of the phase responses is termed the phase matrix. Both matrices are arranged such that increasing temporal modulation frequencies (ω) are shown along the abscissa and increasing spectral modulation frequencies (Ω) are shown on the ordinate. Each matrix has 4 quadrants, where the first quadrant plots the range of spectral and temporal modulations that produced downward sweeping ripples, and the second quadrant plots the same spectral and temporal modulation parameters but for upward sweeping ripples. The third and fourth quadrants provide redundant information for reasons explained below.

The STRF for each neuron was constructed from the combination of ripple stimuli presented, where each ripple was weighted by its magnitude value in the magnitude matrix and temporally shifted by its phase value in the phase matrix. The linear summation of those weighted and phase shifted ripples generated the cell's STRF (Figure 2.1F).

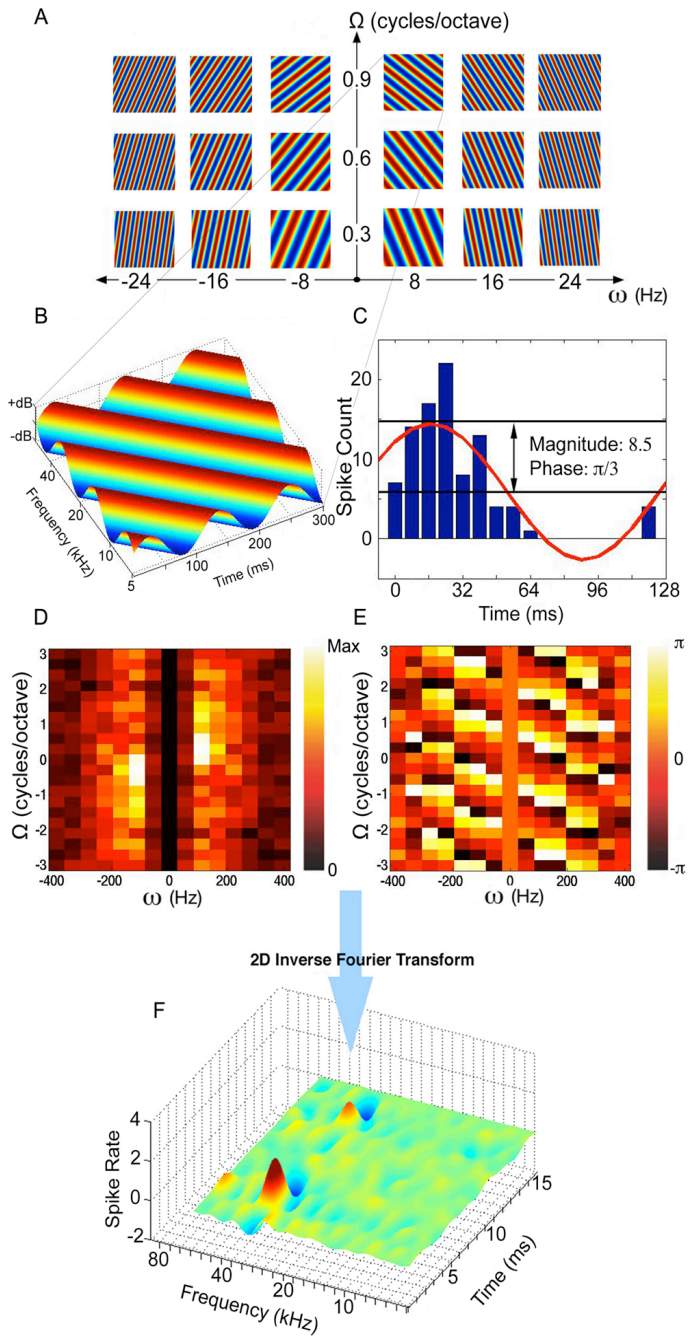

Figure 2.1: Constructing a spectrotemporal receptive field (STRF) from neural responses to moving ripple stimuli. (A) A subset of moving ripple stimuli with different temporal and spectral modulation rates. (B) A single moving ripple stimulus with a temporal modulation rate (ω) of 8 Hz and a spectral modulation rate (Ω) of 0.9 cycles per octave. (C) Neural response of an IC cell to the ripple stimulus shown in (B). Magnitude and phase of the response are extracted by calculating the amplitude and unwrapped phase of the fundamental present in the recorded period histogram with its period equivalent to one ripple cycle (1/ω). (D) Magnitude matrix showing response magnitudes for all ripple stimuli presented. (E) Phase matrix showing response phases to the same ripple set. (F) By using a range of temporal and spectral modulation rates (ω, Ω), a ripple transfer function is obtained with its magnitude (D) and phase (E) derived from the neural responses to each ripple stimulus. An STRF for the neuron is constructed by linearly summing each ripple stimulus in our set, scaled by its response magnitude and shifted by its response phase, which is equivalent to taking the inverse Fourier transform of the ripple transfer function.

The magnitude and phase matrices together constitute the ripple transfer function defined as follows:

$$T(\omega,\Omega) = M(\omega,\Omega) \cdot \exp[j \cdot \Phi(\omega,\Omega)] \quad (2.2)$$

where M is the response magnitude, Φ is the phase of the response and j = $\sqrt{-1}$.

It is important to note here the complex conjugate symmetry in the ripple transfer function, where:

$$T(-\omega,-\Omega) = T^*(\omega,\Omega) \quad (2.3)$$

This symmetry enabled us to measure the transfer function in only two quadrants of ripple space (first and fourth), where positive and negative values were used for Ω and only positive values were used for ω. Using Equation 2.3 enabled us to deduce the transfer function in the other two quadrants (second and third). A value of zero was assumed for ripples with no temporal modulation ($\omega = 0$).

The STRF of the cell was constructed by taking the inverse Fourier transform of the ripple transfer function in two dimensions:

$$strf(t,x) = F^{-1}_{t,-x}\{T(\omega,\Omega)\} \quad (2.4)$$

where $F^{-1}\{.\}$ designates the inverse Fourier transform with respect to ω and Ω.

As seen in Figure 2.1F, the STRF shows the frequency tuning of the neuron on the logarithmic spectral axis. The temporal axis reflects the impulse response of the cell.

The significant portions of each STRF were obtained by estimating response variability using the bootstrap method (Efron and Tibshirani 1993). By resampling with replacement, an estimate of the variability of each point on the STRF was obtained

(Depireux, Simon et al. 2001). We used the same criteria of Fritz, Shamma et al. (2003) for significance. Regions of the STRF with values that exceeded ±2.5 standard deviations from the mean were considered significant whereas other regions were regarded as measurement noise.

Predicted Call Responses from Convolutions

After obtaining an STRF for an auditory neuron, the response of the cell to any given stimulus can be predicted by performing a convolution in the temporal dimension and integration in the spectral dimension with the spectrogram of the stimulus. In the discrete case, the predicted response is the product of a temporally reversed STRF with the spectrogram of the stimulus summed over frequency and time.

$$r(t,x) = \sum_{x} \int S(\tau,x) \cdot strf(t-\tau)\partial\tau \qquad (2.5)$$

The predicted response, r, was then rectified by setting any negative predicted response value to zero, $r = max(0,r)$. This rectification was performed in order to conduct an accurate correlation between the predicted response and the actual response of the cell and to compensate for the half-wave rectification non-linearity present in extracellular recordings of auditory neurons.

STRFs of IC neurons were convolved with each of the communication calls and a predicted response for each call was calculated using Equation 2.5. The highest magnitude of predicted response across all calls was then scaled to the highest magnitude of actual response across the same set of calls. The scaling allowed us to analyze the relative magnitudes of the predicted responses and compare each with the response

magnitudes that were actually evoked by each call. A correlation coefficient was then calculated for each call to assess the accuracy of the predictions. Calls that had weak or no responses in both the predictions and in the actual responses of the cell produced low correlation values due to noise in the STRF and the cell's weak response. These calls were used to demonstrate that STRFs not only predicted calls to which the neurons responded but also predicted the calls to which the neurons failed to respond. No correlation coefficient was calculated for calls that evoked weak responses and for which weak responses were predicted by convolution.

Inseparability and Direction-Selectivity Analysis

To assess the degree of STRF inseparability, we performed singular value decomposition (SVD) (Depireux, Simon et al. 2001). SVD can be viewed as decomposing the STRF into a linear sum of separable matrices, with each matrix weighted by its singular value on how much it can estimate the STRF as a product of two vectors as in $strf = \sum \lambda_i u_i \cdot v_i^T$ where T denotes the Hermitian transpose and λ's being the singular values with $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \ldots$ Therefore, a separable STRF will have a first singular value (λ_1) that dominates all the other singular values. The more dominant is the first singular value compared to the other singular values, the more separable is the STRF. With this in mind, we defined an inseparability index (*Ins*) as follows:

$$Ins = 1 - (\lambda_1^2 / \sum_i \lambda_i^2) \qquad (2.6)$$

A value of *Ins* near zero indicates a separable STRF whereas a value closer to 1 shows a high degree of inseparability.

To specifically investigate the direction selectivity of STRFs and its contribution to inseparability, we calculated a direction selective index for each STRF. The index was calculated from the total power in the first quadrant of the magnitude matrix, P_1, compared with the total power in the second quadrant, P_2. If the STRF favors one ripple direction over the other, more power will be present in the corresponding quadrant of the transfer function versus the other, i.e. an STRF that favors downward moving ripples will have more power in the first quadrant because the first quadrant of the magnitude matrix is composed of the response magnitudes evoked by downward sweeping ripples. The direction selective index (*DSI*) is therefore defined as follows:

$$DSI = (P_2 - P_1)/(P_1 + P_2) \qquad (2.7)$$

where total power is computed by adding the square of the magnitudes in each quadrant. A negative value of *DSI* will indicate direction-selectivity for the downward direction; a positive value shows upward selectivity and no direction-selectivity for values near zero. Note that relative power was examined in only two quadrants of the transfer function due to the complex conjugate symmetry mentioned in Equation 2.3.

Velocity Tuning

We analyzed the tuning of IC neurons to FM velocity in a similar manner that had been applied to receptive fields of MT neurons and their selectivity for visual speed (Perrone and Thiele 2001). Since the velocity of a moving ripple is expressed as the ratio

of temporal to spectral modulation rates ($v = \omega/\Omega$), the magnitude matrix could be then divided into isovelocity lines passing through the origin where FM velocity is directly related to the slope of each line. Therefore, an IC neuron that is tuned for a particular FM velocity will have its ripple responses both elongated and oriented along a specific velocity-line in the magnitude matrix, where the slope of that line indicates the optimal velocity of the neuron.

To quantify the orientation of ripple responses in one quadrant of the magnitude matrix, the responses were normalized and then fit by a rotated two-dimensional Gaussian as follows:

$$G(\omega,\Omega) = \exp(-(\omega')^2 / s_x^2) \times \exp(-(\Omega')^2 / s_y^2) + p \qquad (2.8)$$

where $\omega' = (\omega - x)\cdot cos(\theta) + (\Omega - y)\cdot sin(\theta)$ and $\Omega' = -(\omega - x)\cdot sin(\theta) + (\Omega - y)\cdot cos(\theta)$ are the rotated coordinates of the Gaussian around its peak (x,y) by the angle θ. The spread of the Gaussian along the rotated dimensions (ω', Ω') is indicated by s_x and s_y respectively. A constant pedestal value (p) is added and then G is normalized relative to its maximum. The parameters x, y, s_x, s_y, θ and p were nonlinearly optimized using the Nelder-Mead direct search method (fminsearch in MatLab) to minimize the sum of the squared deviations, or mean-squared error (E), between the neural responses and G values as in:

$$E = \frac{1}{N}\sum_j (R_j - G_j)^2 \qquad (2.9)$$

where R_j is the response magnitude at temporal and spectral modulation rates (ω_j, Ω_j) and G_j is the Gaussian fit value at the same modulation rates. To evaluate the goodness of the

fit, the responses of each neuron were correlated with the Gaussian model and neurons with a correlation coefficient less than 0.7 were not considered.

Velocity tuning of IC neurons was then assessed by comparing the orientation of the Gaussian (θ) to the slope of the velocity-line that passes through the peak response at the (*x*,*y*) location. We define the orientation error (OE) of a neuron as the difference between the angle of optimal velocity tuning (β) and the orientation angle (θ), OE = β – θ, where β = *atan(y/x)* and both angles are in reference to the vertical axis. Therefore, if the Gaussian fit had its main axis oriented towards the origin (0,0) then it would have an OE of zero, which suggests a strong tuning for FM velocity. Alternatively, higher absolute values of OE would indicate lesser degrees of velocity tuning.

Decomposition of Conspecific Vocalizations

Since moving ripples make up the Fourier basis components of the spectral-temporal domain, each syllable of a bat call was decomposed into a linear sum of moving ripples as described by Singh and Theunissen (2003). This was accomplished by performing a windowed Fourier transform in two-dimensions on the spectrogram of that syllable after subtracting its mean as in:

$$T_c(\omega,\Omega) = F_{t,x}\{S(t,x) - \bar{S}\} \qquad (2.10)$$

where $F_{t,x}$ *{.}* denotes the Fourier transform with respect to time and logarithmic frequency. S is the spectrogram of a bat call and $\bar{S}$ is its mean over time and frequency (DC component). A magnitude matrix for each syllable was constructed by plotting the magnitude of each ripple present in the call with a range of temporal modulation rates

from –400 to 400 Hz and a range of spectral modulation rates from –5 to 5 cycles/octave. Similar to the magnitude matrix of an IC neuron showing the magnitude of each ripple present in the STRF, the magnitude matrix of a bat call shows the magnitude of the spectral and temporal modulations (ripples) present in each call. The average FM velocity present in each syllable of a bat vocalization was evaluated by calculating the slope of the velocity-line passing through the origin and the peak value of its magnitude matrix.

Surgical procedures

Surgical and pharmacological procedures, electronic equipment, sound generation, and criteria for isolating single neurons were described in previous publications (Klug, Bauer et al. 2002). The bats were prepared under general anesthesia (IsoFlo, Abbott Labs), and lidocaine (Abbott Labs, North Chicago, IL) was applied topically to all open wounds. After a small metal rod was mounted on the skull, a small hole was made in the skull over the IC and the bat was then transferred to a heated recording chamber, where it was placed in a restraining cushion constructed of foam molded to the animal's body. The metal rod was attached to a bar mounted on the stereotaxic instrument. The electrode was positioned over the IC and was subsequently advanced from outside of the experimental chamber with a piezoelectric microdrive (Burleigh 7121W). Recordings were made at depths ranging from ~350 – 1600 μm, which covered most of the dorsoventral extent of the central nucleus of the inferior colliculus. Recordings were begun after the bats recovered from the anesthetic and thus all data were obtained from awake animals. If the bats showed signs of discomfort, doses of the neuroleptic, ketamine hydrochloride (Vetamine, Mallinckrodt Veterinary, 1/40

dilution, 0.01cc injection) were administered. All experimental procedures were in accordance with a protocol approved by the University of Texas Institutional Animal Care and Use Committee.

Electrodes

Single units were recorded either with a single micropipette filled with buffered 1M NaCl or with 5 barrel "piggyback" multibarrel micropipettes (Havey and Caspary, 1980) with a single barrel pipette attached for recording single unit activity. One barrel of the five-barrel pipette was the balancing barrel and was filled with buffered 1M NaCl and the other barrels with bicuculline methiodide (Sigma, St. Louis, MO), an antagonist of $GABA_A$ receptors, and/or with the glycine receptor antagonist, strychnine HCl (both were 10 mM in 0.165 M NaCl, pH 3.0, Sigma, St. Louis, MO). For bicuculline and strychnine, retention currents were negative and ejection currents were positive. The drug and balancing barrels were connected via silver-silver chloride wires to a six-channel microiontophoresis constant current generator (Medical Systems Neurophore BH-2) that was used to generate and monitor ejection and retention currents. The sum channel was employed to balance current in the drug barrels. The recording barrel was connected to a Dagan AC amplifier (model 2400).

RESULTS

This study is based on 114 neurons recorded from the central nucleus of the inferior colliculus (IC) of Mexican free-tailed bats. Similar to other STRF studies in the IC (Poon and Yu 2000; Escabi and Schreiner 2002), cells in our study responded in one of three ways to the moving ripple stimuli. Eight neurons (~7%) failed to respond to any of the ripple stimuli and no STRF could be extracted for these cells. Another 47 neurons (~41%) either responded to only a few of the ripple stimuli or did not phase-lock well to modulations in the stimuli, and thus STRFs could not be derived for these neurons either. However, 59 neurons (~52%) responded to ripple stimuli with phase-locked responses and STRFs were constructed for these neurons. In these 59 neurons, their tone evoked best frequencies (BF), the frequency requiring the lowest intensity to evoke spikes, ranged from 16-49 kHz, although most (30/59, 51%) had BFs between 21-25 kHz. The BFs of 24% (14/59) were between 16-20 kHz and 17% (10/59) were tuned between 26-30 kHz. The remaining 8% were tuned to frequencies above 30 kHz. While most of the neurons in our sample were tuned to a narrow range of frequencies, between 21-25 kHz, this frequency range is overrepresented in the cochlea as well as in other auditory brainstem nuclei of Mexican free-tailed bats (Vater and Siefer 1995; Klug, Bauer et al. 2002; Xie, Meitzen et al. 2005).

A representative STRF is shown in Figure 2.2A and illustrates many of the features common to the STRFs we obtained from the other 58 neurons. The STRF had a clear excitatory region surrounded temporally and spectrally by inhibitory regions. In addition, the neuron had an onset inhibitory region at a frequency corresponding to the

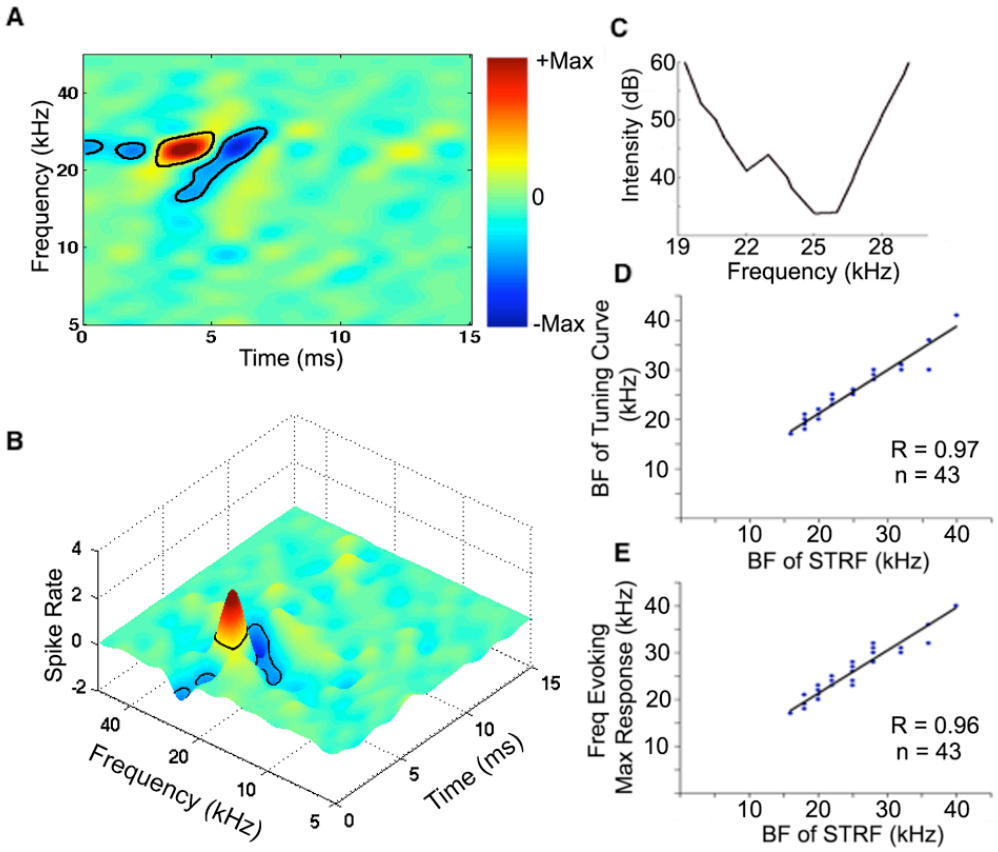

Figure 2.2: Features of spectrotemporal receptive fields (STRF) and their agreement with tone-evoked responses.
(A, B) STRF generated by a family of moving ripples is shown as a 2D plot and the same STRF plotted in 3D is shown below. Excitation is indicated in red and inhibition in blue with black contour lines depicting significant regions (see Methods). The best frequency of the STRF (BF_{STRF}), the frequency on the spectral axis of the STRF that had the highest peak on the temporal axis, was ~25 kHz. The STRF also had both onset and offset inhibitions, whose frequencies corresponded to but were temporally separated from the excitatory region, and an inhibitory region that flanked the low frequency side of the excitatory region. (C) The tuning curve of the same neuron generated by tone-evoked responses. The BF, the frequency to which the neuron was most sensitive, of the tone evoked tuning curve was also 25kHz. (D, E) Plots showing agreement between tone evoked responses and STRFs in 43 IC neurons.

excitatory frequency but with a shorter latency that preceded the excitation in time, and an offset inhibitory region at the same frequency as the excitatory frequency but with a longer latency. Surrounding inhibitory regions on only the low frequency side of the excitatory regions, as occurred in the cell in Figure 2.2A, was seen in ~11% of the neurons. Inhibitory regions on both the high and low frequency side of the excitatory region were seen in the majority of STRFs (~63%), although their relative magnitudes and shapes differed greatly among different cells. In others (~5%), the surround inhibition was only above the excitatory region, while other cells had no surrounding inhibitory regions (~21%). All STRFs also had onset or offset inhibitory regions, or both onset and offset inhibition, whose frequencies corresponded to but were temporally separated from the excitatory regions. The presence of onset and/or offset inhibitory regions is consistent with results reported in previous studies (Pollak and Park 1993; Covey, Kauer et al. 1996).

Receptive Fields and Tuning Curves

To verify the validity of STRF excitatory regions, we compared tone-evoked tuning curves of 43 neurons with their corresponding STRFs. The best frequency of the tuning curve (BF_{TC}) was compared to the best frequency of the STRF (BF_{STRF}), the frequency on the spectral axis of the STRF that had the highest peak on the temporal axis. In almost every neuron, the BF_{TC} was in close agreement with the BF_{STRF} obtained from responses to ripple stimuli (R=0.97, Figure 2.2D). We also compared the frequency of the tone burst that evoked the highest spike-count with the neuron's BF_{STRF} (Figure 2.2E). Here too there was a strong correlation between the two measures (R = 0.96). These

results indicate that the STRFs extracted the excitatory response properties from these IC neurons.

Predicting Neural Response to Conspecific Vocalizations

To further establish the validity of STRFs, we asked whether the STRFs could predict how the neurons would respond to complex signals that were not used to generate their STRFs. To do this, we presented a suite of 6 conspecific communication calls to 44 neurons that were held long enough so that we obtained both their STRFs and responses to the calls. We compared the responses that were actually evoked by the 6 calls to the predicted responses for those calls that were generated by convolving the neuron's STRF with the spectrogram of each communication call. We then evaluated: 1) the degree to which the temporal response patterns predicted by the convolutions matched the temporal response patterns that were actually evoked by the calls; and 2) whether the convolutions predicted which calls actually evoked responses and which calls failed to evoke responses.

The agreement among the predicted and evoked responses differed among the 44 neurons and ranged from poor to very good. Although the 44 neurons displayed various degrees of predictability, there were no obvious differences in the STRFs of neurons whose STRFs yeilded high predictabilities from those that yielded low predictabilities. In ~27% (12/44) of the cells the median correlation of predicted and evoked responses ranged from 0.1-0.29. However, in 73% (32/44) of the cells the median correlations were at least 0.3 and typically much higher. In these neurons, the STRFs not only predicted

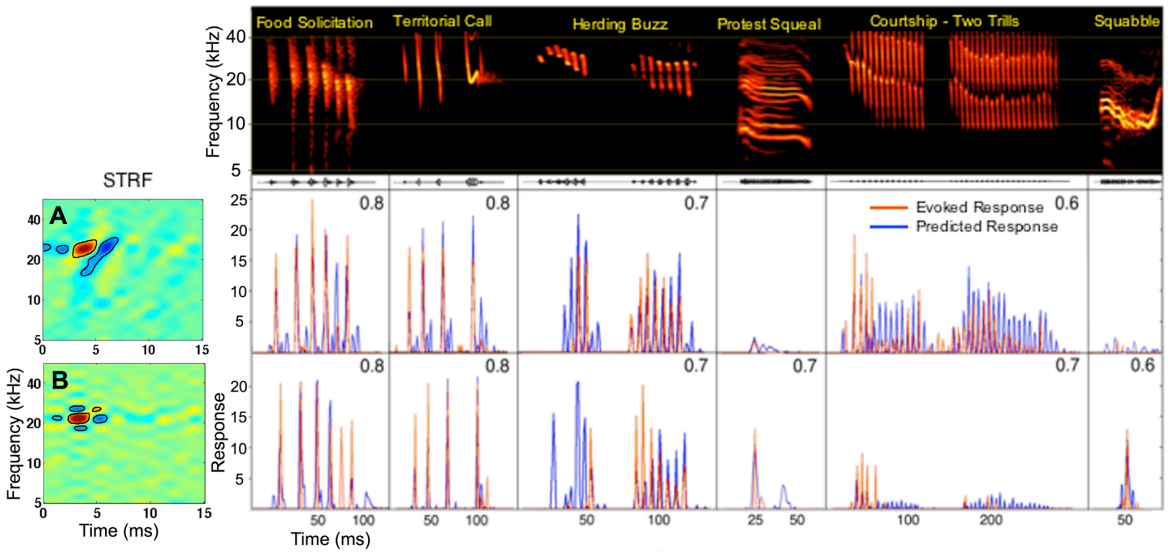

Figure 2.3: STRFs provide accurate predictions of responses to species-specific calls. Spectrograms of each species-specific vocalization are shown in top panel with the evoked response of each IC neuron (red) and the predicted response of its STRF (blue) displayed below each call. Each row shows the predicted and evoked responses of an IC neuron with its STRF on the left. Predicted responses were generated by convolving the neuron's STRF with the spectrogram of each call. Correlation between predicted and actual response is shown in upper right of each panel. (A, B) Two neurons in which there were high correlations between their STRF predictions and their actual responses. Convolutions predicted the call selectivity of these neurons since they predicted very low response magnitudes for the calls that evoked little or virtually no responses, but predicted high response magnitudes for those calls that evoked strong responses. Those calls that evoked little or no responses were used to demonstrate that convolving STRFs with those calls also predicted little or no activity, but no correlation coefficient was computed for those calls.

the calls to which the neurons responded and their temporal discharge patterns, but also predicted the calls to which the neurons failed to respond. Two examples of neurons that had good predictions, are shown in Figure 2.3A,B. Both neurons had similar BFs, 25 kHz (Figure 2.3A) and 23 kHz (Figure 2.3B), and both neurons responded similarly to some calls and differently to others. This shows that even minor differences in the spectrotemporal arrangement of excitatory and inhibitory regions can have a substantial influence on shaping responses to complex signals.

Blocking Inhibitory Receptors

One of the questions we address is whether the inhibitory regions in the STRFs were caused by inhibition at that IC cell or whether the inhibition occurred in a lower nucleus and the spike suppressions were then inherited by the IC cell. These alternatives are not mutually exclusive, and it may be that some inhibitory frequency regions were caused by inhibition at the IC cell whereas other inhibitory regions of the same cell were inherited from lower nuclei.

To evaluate the degree to which inhibition was generated in the IC, we recorded STRFs in 16 cells both before and while inhibition was blocked by the iontophoretic application of bicuculline, an antagonist of $GABA_A$ receptors, and/or strychnine, an antagonist of glycine receptors. In 12 of 16 cells, some of the inhibitory fields in their STRFs were largely eliminated and others were greatly reduced when inhibition was blocked, as illustrated by the neuron in Figure 2.4. Although the blockers eliminated some inhibitory regions in 12/16 neurons, in none of those neurons could we completely

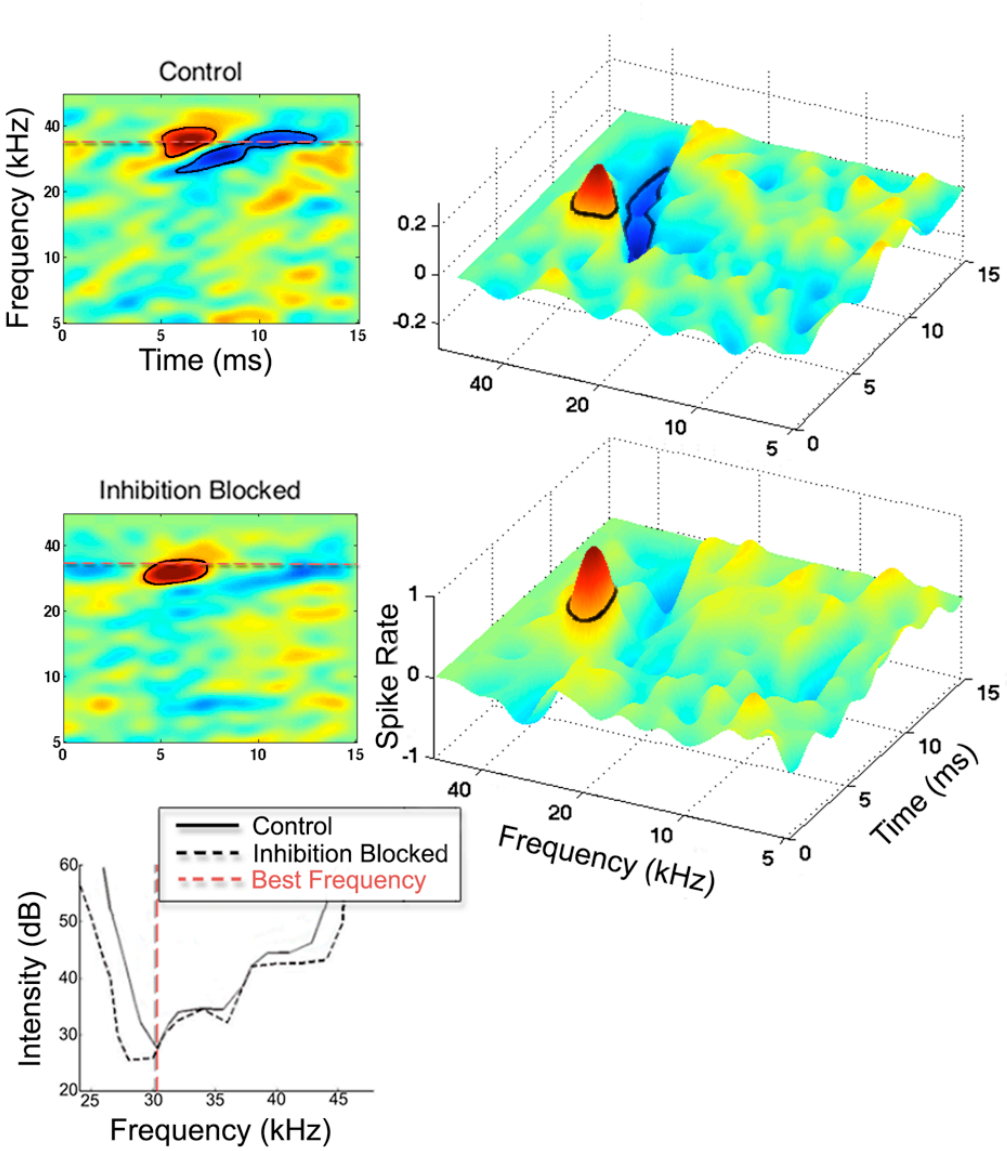

Figure 2.4: Effects of blocking inhibition on the STRF and tuning curve of an IC neuron.

The reduction in surrounding inhibitory regions in the STRF of a neuron before blocking inhibition (control) and while inhibitory receptors were blocked. Blocking inhibition also caused a small expansion of the neuron's excitatory region. Both bicuculline (60nA injection current) and strychnine (60nA injection current) were applied to the cell.

eliminate all inhibition. Presumably the remaining inhibitory regions were either inherited from lower regions or remained because the drugs failed to block all inhibitory receptors on the neurons.

Inseparability and Direction-Selectivity

As shown in the previous sections, and as confirmed by the STRF predictions, the STRFs were able to extract excitatory and inhibitory response properties in many IC neurons. In this section, we analyzed the arrangement of excitatory and inhibitory regions within the STRFs of the 32 neurons whose median predictability for calls was 0.3 or above. Since most bat communication and echolocation calls are composed of FM sweeps, we investigated the selectivity of the 32 neurons for the direction of spectral motion by evaluating the spectrum-time inseparability of their STRFs.

Similar to the analysis used in the visual system to assess directional selectivity of motion (Reid, Soodak et al. 1991 ; Priebe, Cassanello et al. 2003) we calculated an inseparability index (Ins) for each STRF to obtain information about whether the spectrotemporal organization of its receptive fields had features required for FM sweep direction selectivity. An STRF whose spectrotemporal features are fully separable can be expressed as the product of a function in time and another in frequency, and thus the same temporal response pattern would be evoked by any excitatory frequency (Sen, Theunissen et al. 2001). Neurons with fully separable STRFs cannot be directionally selective. In order for a neuron to respond more favorably to one FM sweep direction over the other, its receptive field must have different temporal response properties (latencies) for different frequencies, and thus its response to any auditory stimulus has to

take both the spectral and temporal properties of the stimulus into consideration, i.e., its STRF is spectrotemporally inseparable.

To illustrate how these features contribute to directional selectivity, either excitatory or inhibitory regions or both regions in the STRFs of inseparable neurons are tilted in the frequency-time plane (e.g., the STRF in Figure 2.3A). The spectrotemporal tilts represent shifts in latency of either excitation or inhibition along the frequency axis. The tilt in the excitatory field could impart a limited degree of directional preference. The reason is that any point in time an FM signal sweeping in the non-preferred direction will encroach upon a smaller portion of the excitatory receptive field than will a signal that sweeps in the preferred direction. Whether or not tilts occur in the excitatory field, tilted inhibitory fields enhance directional preferences, or even create them, because signals sweeping in the non-preferred direction simultaneously evoke both excitation and inhibition, thereby suppressing responses to that FM direction, whereas signals sweeping in the preferred direction activate excitation and inhibition at different times, thereby allowing the cell to respond to the preferred direction.

To evaluate the degree of inseparability, we calculated the overall inseparability of the STRFs of each of the 32 neurons whose median predictability for calls was 0.3 or above. Inseparablility is not an all or none property and STRFs can be partially inseparable. Most IC neurons had some degree of inseparability, with a mean inseparability index of 0.33 ± 0.14. The distribution of inseparability indices is shown in Figure 2.5A.

To evaluate the relationship between inseparability and FM directional selectivity, we determined the preference for FM sweep direction in the 32 neurons and correlated their directional selectivity with their inseparability index. We evaluate preferences for FM sweep direction by referring to each neuron's magnitude matrix that was used to construct its STRF. As described in the Methods Section, each moving ripple we presented was constructed from a particular rate of spectral modulation (SMR) and a rate of temporal modulation (TMR), which in combination imparted a frequency modulation sweep with a certain velocity and upward or downward direction (Figure 2.1B). The vigor with which a neuron responded to each downward and upward moving ripple is documented in its magnitude matrix (Figs. 1D, 6A). Since ripples in the first quadrant of the magnitude matrix were all downward sweeping, and ripples in the second quadrant were all upward sweeping, we calculated a direction selectivity index (DSI) for each neuron by comparing the relative power between the first and second quadrants in the magnitude matrix of each neuron**.** Negative values show that downward sweeping ripples in the first quadrant of the matrix evoked the strongest responses and positive values indicate strongest responses to upward sweeping ripples in the second quadrant. We considered neurons to be directional selective if their absolute DSI was greater than 0.25. As seen in Figure 2.5B, the DSIs of most (22/32) neurons were more negative than -0.25, showing that most IC neurons favored downward sweeping ripples. The DSIs of only 2 neurons were greater than +0.25 and were selective for upward sweeps. The remaining 8 neurons had DSIs between $\pm$ 0.25 and were considered non-selective. The average direction selectivity index was –0.37 $\pm$ 0.3.

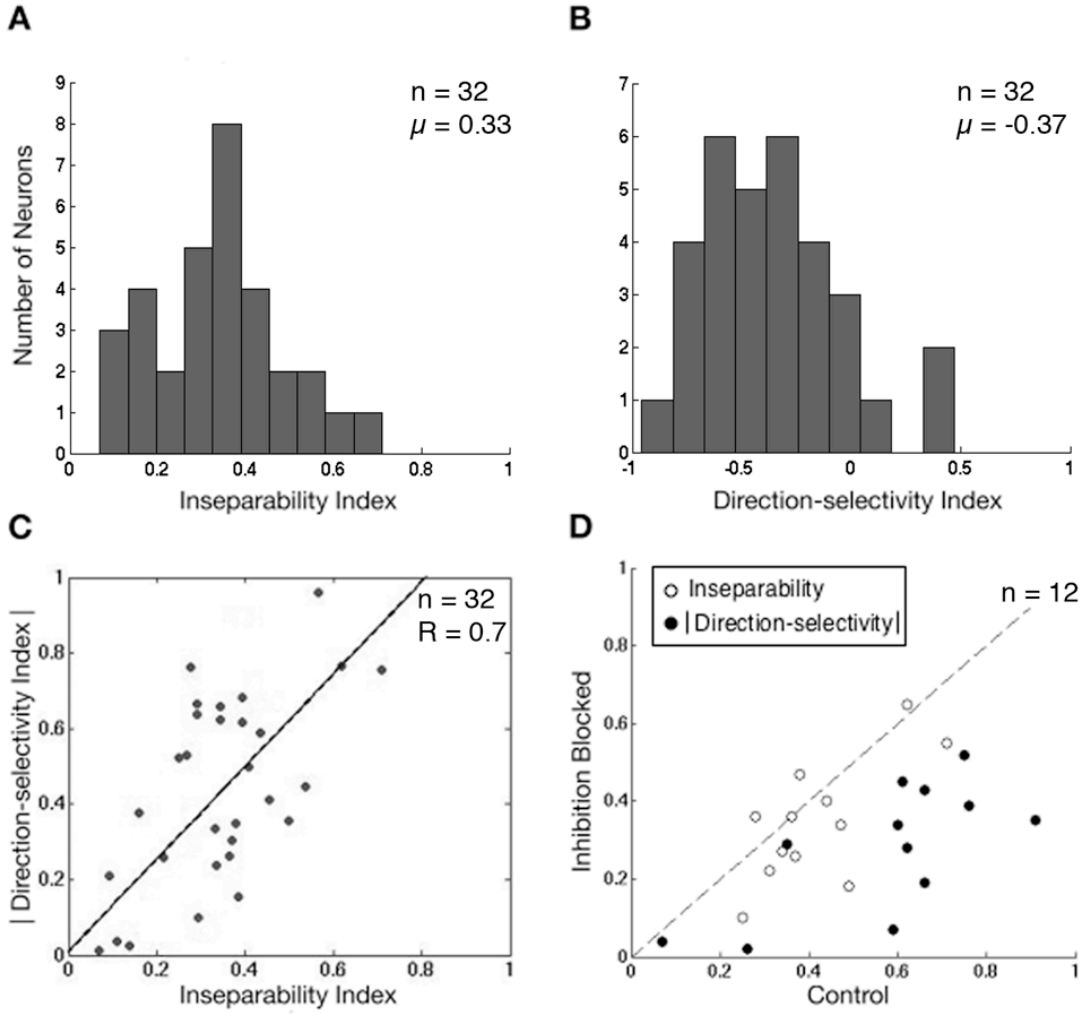

Figure 2.5: Inseparability and direction-selectivity of STRFs.
(A) Distribution of the inseparability indices across all IC neurons with a valid STRF that yielded a predictability of 0.3 or higher. (B) Distribution of direction-selectivity indices across the same neurons. (C) Correlation between the two indices indicating a strong contribution of inseparability of the STRFs to direction-selectivity. (D) Blocking inhibitory receptors of 12 IC neurons reduced their direction-selectivity and to a lesser extent their spectrum-time inseparability, indicating that both properties are shaped by inhibition.

To examine the relationship between inseparability and direction-selectivity, we plotted the absolute magnitude of the direction selectivity index ($|DSI|$) against the inseparability index for each of the 32 neurons. As shown in Figure 2.5C, the inseparability index was strongly correlated with the direction selectivity index ($R = 0.7$). This suggests that spectrum-time inseparability of IC neurons is the main contributor to the direction-selectivity. Moreover, since most direction selectivity indices were negative, IC neurons in our study had a preference for responding to downward compared with upward sweeping ripples, and presumably other FM signals as well.

Inhibition Shapes Directional Selectivity

Previously we showed that blocking inhibition eliminated most of the inhibitory fields in the STRFs of the majority of neurons. If the structure (tilting) of surround inhibition in the STRF contributes to inseparability, then blocking inhibition should not only reduce inseparability but also should reduce direction-selectivity. To show this relationship, we calculated inseparability and direction selectivity indices for the 12 neurons whose STRFs were changed when inhibition was blocked. In the majority of these neurons, blocking inhibition caused a reduction in both inseparability and direction selectivity indices (Figure 2.5D). An example showing the change in a neuron's magnitude matrix of responses evoked by upward and downward sweeping ripples before and while inhibition was blocked is shown in Figure 2.6. Blocking inhibition caused a 56% reduction in direction selectivity indices, declining from a mean of –0.5 to –0.24, whereas inseparability decreased by 20%, from a mean of 0.38 to 0.30.

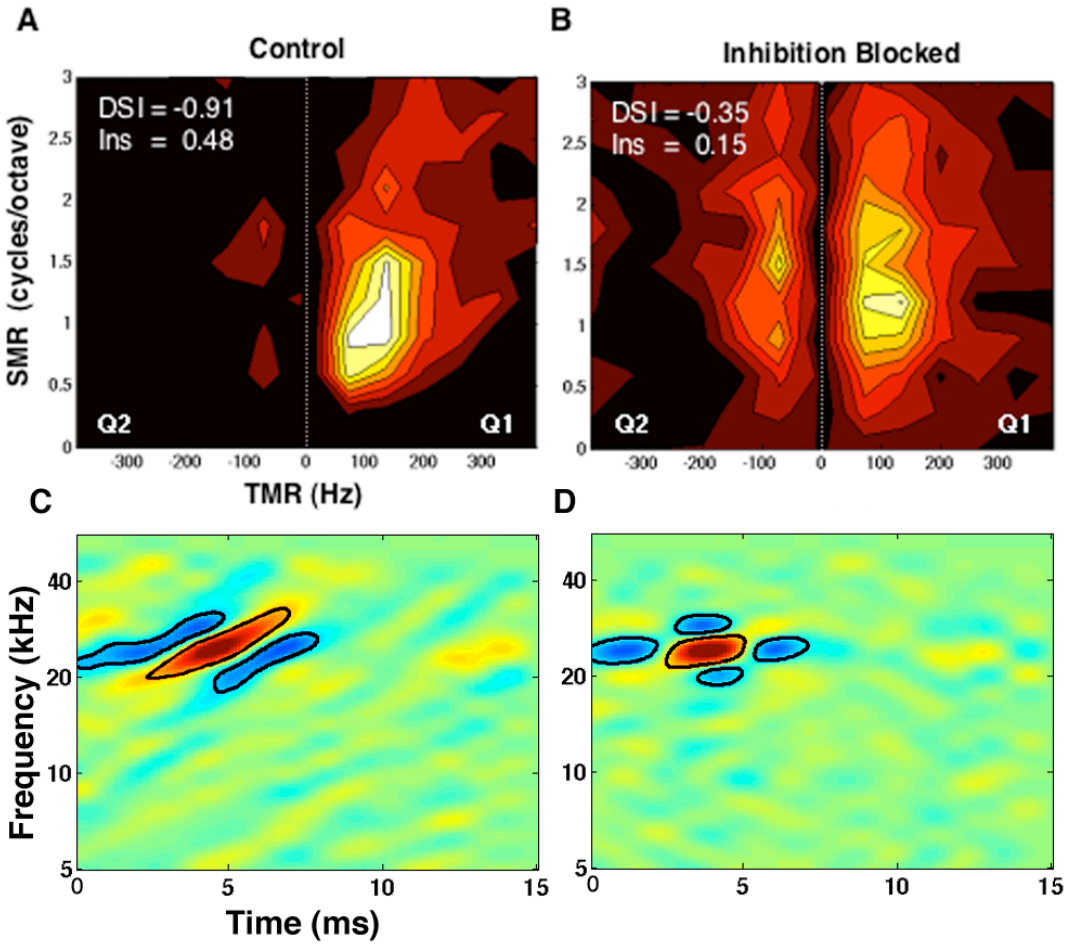

Figure 2.6: Effects of blocking inhibition on neural responses to moving ripples. Contour plots of the magnitude matrices of one neuron containing the ripple responses obtained before (A) and while inhibition was blocked (B). Before inhibition was blocked, the neuron responded strongly to downward sweeping ripples in quadrant 1 and hardly at all to upward sweeping ripples in quadrant 2, and thus was directionally selective. Its directional selectivity index was -0.91. When inhibition was blocked, the range of spectral modulation rates (SMR) as well as temporal modulation rates (TMR) to which the cell responded increased in both quadrants of the magnitude matrix. Thus the directional selectivity index was reduced to -0.35 due to the smaller difference in overall power between the two quadrants. The resulting STRFs before and while inhibition was blocked are shown in (C) and (D) respectively.

One explanation for this difference in the effects of inhibition on direction selectivity and inseparability is that blocking inhibition frequently causes an expansion of the excitatory receptive field, where the expansion is not uniform along the spectrotemporal axis (Yang, Pollak et al. 1992; Palombi and Caspary 1996; LeBeau, Malmierca et al. 2001; Klug, Bauer et al. 2002). Such a non-uniform expansion of the excitatory receptive field contributes to inseparability. Thus, blocking inhibition could have had two opposing influences on inseparability. It reduced inseparability due to the reduction or elimination of tilted inhibitory regions thereby reducing or eliminating the power asymmetry between the two quadrants of the magnitude matrix. However, a non-uniform expansion in excitation could enhance inseparability if responses to spectral and temporal modulation rates were different for the preferred compared with the non-preferred direction (Figure 2.6B), thereby creating spectrotemporal asymmetries between the two quadrants of the magnitude matrix. The net effect should be a substantial reduction of direction selectivity coupled with a smaller change in inseparability, which is the result we obtained.

Comparing Properties of Bat Calls with Response Features of IC Neurons

In the section above we showed that most IC neurons in the Mexican free-tailed bat were directionally selective, favoring the downward FM direction. This directional preference agrees with the prominence of downward FM components present in their echolocation and social communication calls, an agreement that can be qualitatively assessed by visually inspecting their spectrograms (Figure 2.3). Here we evaluate the

velocities of the FM sweep components in their calls and compare them to the FM velocities of the ripple stimuli to which IC neurons responded most strongly.

We decomposed the spectrogram of each call into a set of ripples with different magnitudes and phases, as described in the Methods Section. This allowed us to compare the magnitudes of the particular ripples present in each call to the magnitude of the response evoked by each of the ripples we presented, as revealed in the neuron's magnitude matrix. Additionally, decomposing a call into its ripple components provides information about both its FM direction and its sweep rate (FM velocity). To further evaluate whether acoustic features of their natural calls correspond to neuronal response features, we calculated the FM velocities to which IC neurons were tuned and compared them to the FM velocities of their natural calls.

Before turning to the ripple decomposition of the natural calls and how that decomposition was compared to the ripple composition of the neuron's STRF, we first illustrate how the decomposition was used to assess FM velocity by considering a simpler signal, a synthetic downward FM sweep. As shown in Figure 2.7, the ripple composition of an FM sweep is aligned diagonally on its magnitude matrix and clusters around a line that passes through the origin (0,0). The slope of the line indicates the FM velocity (sweep rate) of the signal. Since the velocity of a moving ripple is equal to the ratio of its temporal to spectral modulation rates (v = TMR/SMR), each of the ripples that lie on that line has the same velocity, which is the velocity of the FM sweep.

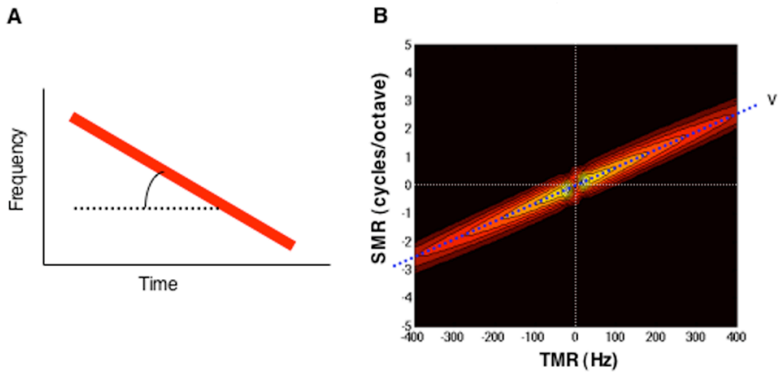

Figure 2.7: Decomposing a synthetic FM sweep into its ripple components.
(A) A spectrogram of a synthetic, downward FM sweep moving at a constant velocity indicated by its slope. (B) Magnitude matrix of the FM sweep shown in (A) obtained by the two-dimensional Fourier transform of its spectrogram. Since FM velocity is equal to the ratio of temporal to spectral modulation rates (TMR/SMR), the ripple composition of the sweep clusters around a line passing through the origin (0,0) with a slope that indicates the FM velocity of the sweep.

If an IC neuron were selective for an FM velocity, in addition to FM direction, then its magnitude matrix, which shows the response magnitudes evoked by each moving ripple, should express similar features to those of the magnitude matrix of a decomposed FM signal. Specifically, those ripples that evoked responses should roughly cluster around a line that passes through the origin of the magnitude matrix, similar to that shown for the synthetic FM sweep. The slope of the line indicates the velocity to which the neuron is best tuned, and most ripple responses should be aligned or oriented along that velocity line. Velocity-selectivity thus requires an orientation or tilt of responses in one quadrant of the magnitude matrix. Therefore, velocity tuning entails inseparability within a particular FM direction (quadrant) in addition to full spectrum-time inseparability (Depireux, Simon et al. 2001). While overall inseparability provides the

substrate for direction-selectivity as discussed above, quadrant (directional) inseparability enables an IC neuron to be selective for the velocity of spectral motion.

To evaluate whether IC neurons were tuned to velocity, we first calculated the quadrant inseparability for the downward direction in the 30 neurons whose STRFs yielded good predictions and had strong responses in the first quadrant of their magnitude matrix. Most neurons in our sample were indeed quadrant inseparable with a mean index of 0.25 ± 0.12. Next we evaluated the degree to which responses were both elongated and oriented along a particular velocity-line in the magnitude matrix. This was computed by fitting the responses of a neuron with a rotated two-dimensional Gaussian (Figure 2.8A, B), which estimated the orientation angle as well as the spread of responses across the magnitude matrix (see Methods). To assess the goodness of the Gaussian fits, we correlated the responses with the resulting fitted model. All 30 neurons had correlation coefficients of at least 0.7, with a mean correlation coefficient of 0.86±0.07. To evaluate the degree to which the orientation of each Gaussian contributed to the goodness of each fit, the fitting was repeated while constraining the orientation angle to zero ($\theta = 0^{o}$). The non-oriented fits decreased the correlation coefficient in most neurons to a mean of 0.78±0.08. Comparing the correlations of the oriented to those of the non-oriented fits suggests that the orientation of the Gaussians contributed significantly ($t = 4.34$, $df = 58$, $p < 0.005$) to the goodness of the fitted model.

The orientation of each Gaussian was then compared with the slope of the velocity line passing through its peak (Figure 2.8B). We refer to this line as the best

velocity (BV) line, which represents an ideal or perfect velocity tuning. If all the responses in the matrix fell on that line and only on that line, the neuron would have fired only to one FM velocity, its BV, and thus would have been perfectly tuned for velocity. Therefore, the deviation of the orientation angle from the angle of the BV-line, referred to as the orientation error, indicates the degree to which an IC neuron is tuned for velocity, where the larger the error the poorer the tuning.

Figure 2.8D plots the orientation errors in the 30 IC neurons. The orientation errors were normally distributed around a mean of zero degrees with a standard deviation of 7 degrees. Additionally, comparing the spread of responses along their orientation angle (s_x) with their spread perpendicular to that angle (s_y) indicates that responses were elongated along their orientation with a median ratio (s_x/s_y) of 5. Their mean ratio of 20 ± 18 was significantly larger than 1, the ratio of a circular shape ($t = 4$, $df = 29$, $p < 0.001$). This shows that most IC neurons in our sample had a strong degree of velocity tuning and that the neurons were tuned to the sweep rate shown by the BV-line. The distribution of the FM velocities to which the neurons were best tuned is displayed in Figure 2.8C. Most IC neurons had a BV between 0-100 octaves/s with a mean of ~60 oct/s, yet some had BVs as high as ~230 oct/s.

Since the majority of IC neurons in our sample had some degree of orientation error, we further examined whether the velocity to which each IC cell is tuned remains constant with different spectral modulation rates. Figure 2.9a shows the velocity tuning of an IC neuron in response to moving ripples with a fixed spectral modulation rate but a

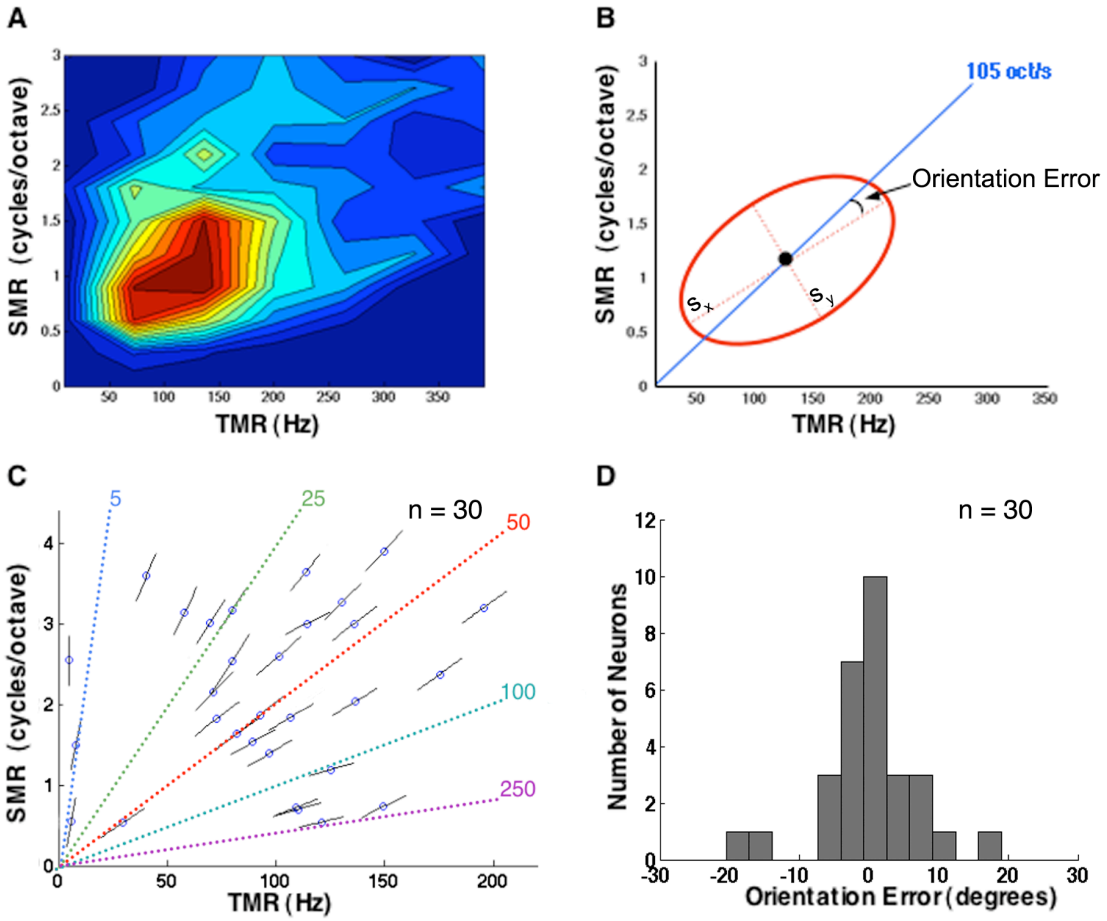

Figure 2.8: Velocity tuning of IC neurons.
(A) Contour plot of the first quadrant in the magnitude matrix of an IC neuron showing responses to downward moving ripples. (B) Responses were fit with a two-dimensional Gaussian, which estimated the orientation angle as well as the spread of responses across the quadrant. The orientation of responses was compared with the velocity line (blue) passing through the peak (black dot). The slope of the line represents the best velocity (BV) of the neuron. The deviation of the orientation angle from the angle of the BV-line (orientation error) indicates the degree to which an IC neuron is tuned for velocity, where the larger the error the poorer the tuning. The neuron shown had a BV of 105 oct/s, an s_x/s_y ratio of 4.5 and an orientation error of 5°. (C) Response peaks and orientation lines are shown for 30 IC neurons in reference to different velocity lines. Most IC neurons had a BV between 0-100 octaves/s, with a mean of ~60 oct/s. (D) Distribution of orientation errors obtained from the 30 neurons showing a mean of 0º and a standard deviation of 7º. This shows that most IC neurons in our sample had a strong degree of velocity tuning.

variable temporal modulation rate and thus a variable FM velocity. Fitting a Gaussian to the responses shows that the peak velocity to which the neuron is tuned does not change with different spectral modulation rates (Figure 2.9b). Examining our population of IC neurons shows that velocity tuning is independent of spectral modulation rates in the majority of these cells (Figure 2.9c). This further validates that IC neurons are strongly tuned for FM velocity.

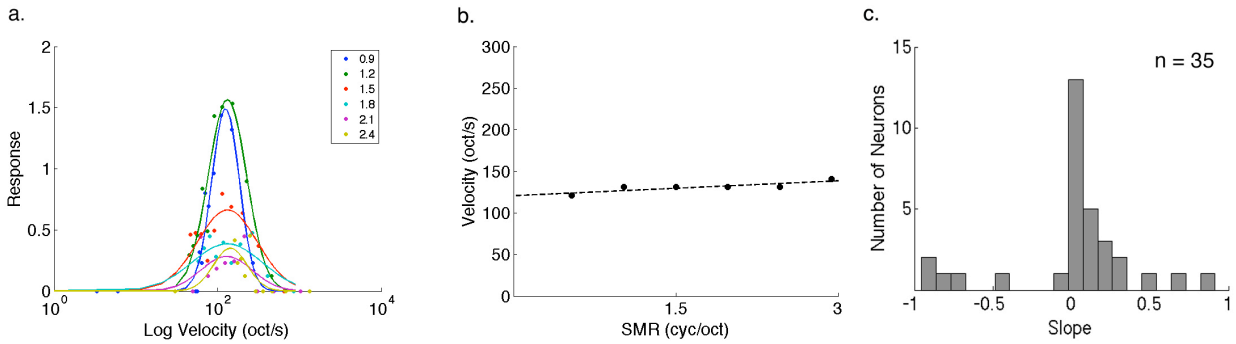

Figure 2.9: Velocity tuning is independent of spectral modulation rate.
(a) Responses to moving ripples with a fixed spectral modulation rate (SMR), as indicated in the inset for each color, and a variable temporal modulation rate (TMR) giving different velocities. FM velocity is plotted on a log scale and responses were fitted with a Gaussian. Note that the peak velocity to which the neuron is tuned remained around 120 oct/s for all SMRs. (b) Plotting BV for the same neuron as a function of SMR shows a flat line with a slope of 0.05. (c) The slope of the linear fit between velocity and SMR for our population of neurons indicated that most neurons had a slope around zero and thus were velocity tuned.

To compare the velocity tuning of neurons with the velocity features of natural calls, we decomposed 21 bat communication and echolocation calls containing 32 syllables or contiguous segments into their ripple components. Figure 2.10 shows spectrograms of four of these syllables with the magnitude matrices of their ripple components. Since

most of these syllables were downward FM sweeps, their ripple spectrum was oriented along the velocity-line corresponding to the average FM velocity present in each syllable. While some calls such as the protest squeal had zero velocity, others reached sweep rates of ~250 octaves/s. The distribution of velocities derived from decomposing 32 syllables are plotted in Figure 2.11A. Finally, we evaluated the degree to which the velocities of the calls that evoked the highest spike rates correlated with the neuron's best velocity for the 17 cells in which median predictions were 0.5 or higher. Fifteen of the 17 neurons responded best to calls that had comparable velocities. Thus, 3 neurons tuned to velocities of 100 oct/sec or higher responded best to calls with velocities of at least 100 oct/sec, while 12 neurons tuned to velocities below 100 oct/sec responded best to calls with velocities below 100 oct/sec. In only 2 neurons was there a mismatch between the neuron's best velocity and the velocity of the call that evoked the maximum response. For the 17 neurons, the correlation between call velocity and tuned velocity was 0.55.

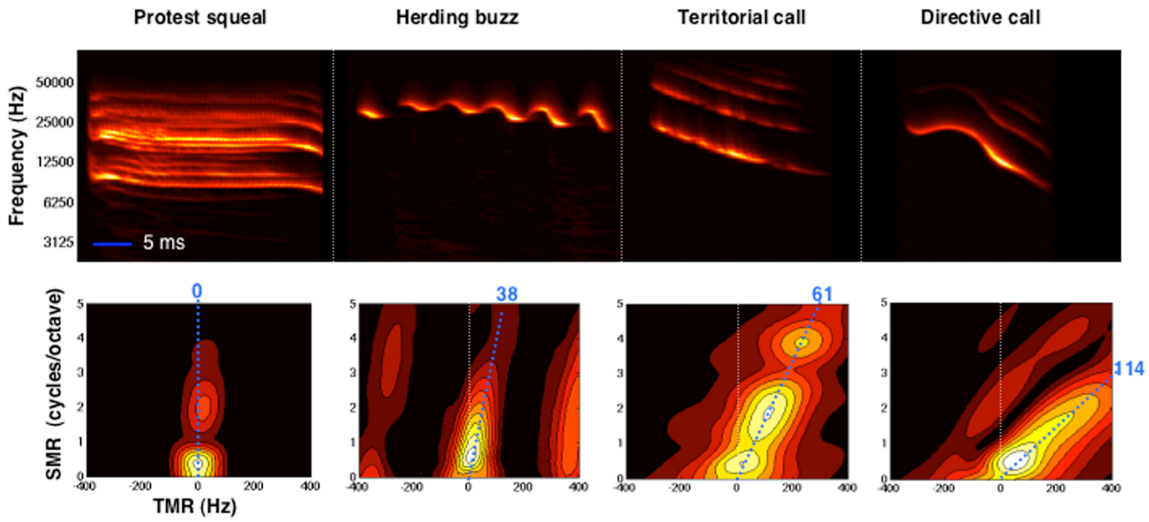

Figure 2.10: Decomposing conspecific calls into their ripple components.
Top row shows examples of four syllables from four bat calls. Bottom row displays their magnitude matrices that show the ripple composition of each syllable. Similar to the decomposition of the synthetic FM sweep, the tilt in each the ripple composition of each matrix indicates the averaged FM velocity present in each syllable. Each syllable's FM velocity is indicated by the dotted blue line in its matrix.

What is apparent is that the FM velocities to which IC neurons are tuned and the velocities present in their conspecific calls correspond closely (Figure 2.11). These results, together with the results shown in the previous sections, suggest that the population of IC neurons in Mexican free-tailed bats are tuned to both the FM direction and the velocities found in their natural signals.

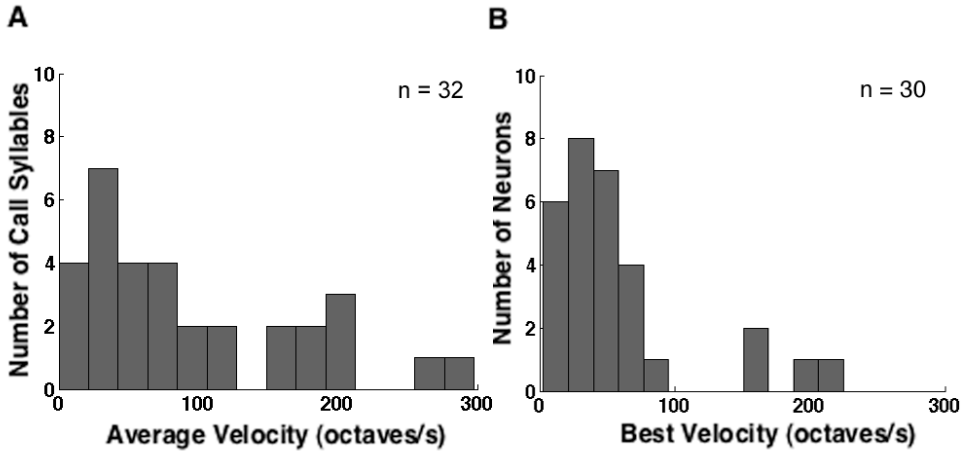

Figure 2.11: FM velocities in communication calls and IC neurons.
(A) Distribution of FM velocities found in 21 calls containing 32 syllables. (B) Distribution of the best velocities to which IC neurons are tuned. The two distributions are well correlated (R = 0.7), suggesting that the FM velocities to which IC neurons are tuned and the velocities present in their conspecific calls correspond closely.

DISCUSSION

Our motivation for generating STRFs was to determine whether we could obtain a description of the spectrotemporal features of excitation and inhibition that is sufficiently general so that we could predict responses to other stimuli and thereby provide an explanation for the behavior of IC neurons. The test we used to evaluate the quality of the description was the degree to which the STRF predicted responses to natural calls. By presenting a family of moving ripples to IC neurons in Mexican free tailed bats, we obtained STRFs from more than half of the neurons in our sample and showed that the STRFs in most of those neurons provide an accurate description of their linear receptive fields. That STRFs captured the excitatory response properties of IC neurons is shown by the high correlations of the BFs evoked by tones and the BF of the STRF, as well as the correlations of the frequencies evoking the highest discharges and the BF of the STRF. The STRFs not only extracted excitatory properties, but also the inhibitory features as shown by the moderately high or very high correlations among predicted and evoked responses in a substantial number of cells. The processing in these neurons was largely linear and thus their STRFs extracted most of their functional properties. The validity of their STRFs was further illustrated by their ability to predict response failures to some calls.

Although the STRFs described the spectrotemporal properties of inhibition and excitation for most IC neurons, in some neurons the convolutions did not provide reliable predictions of the responses that were actually evoked by species-specific calls. The question is why the STRFs failed to predict responses to natural calls in those neurons, or

in other neurons why the family of moving ripples failed to generate an interpretable STRF? One explanation for neurons in which the STRFs provided poor predictions is that they apparently had either static or dynamic non-linear response properties that were stronger than the linear response properties extracted by the STRFs generated by moving ripples. Another explanation is that different stimuli may have generated stronger STRFs or different STRFs than did the family of moving ripples that we employed. Recent studies in both cats (Escabi and Schreiner 2002; Escabi, Miller et al. 2003; Escabi and Read 2005) and songbirds (Woolley, Fremouw et al. 2005) have shown that stimuli with statistics that most closely correspond to natural stimuli generated either stronger or even different STRFs than other stimuli. For example, a recent study by Woolley et al., (2006) of the midbrain of zebra finches showed that STRFs generated with zebra finch songs yielded different STRFs than those generated by modulation limited noise. Moreover, the STRFs generated by song components provided better predictions of the actual responses evoked by song than did the STRFs generated by noise. Following from this, it seems possible that had we used a variety of natural calls, rather than ripples, we might have generated STRFs in a larger number of cells, and in those cells in which the ripple generated STRFs provided poor predictions, somewhat different STRFs would have been generated that may have yielded more accurate predictions of responses to the natural calls.

STRFs Reveal Spectral Motion Detecting Properties of Auditory Neurons

Species-specific calls, or any other signal, can be decomposed into their ripple components and thus the degree to which the ripple composition of any signal matches

those in the neuron's magnitude matrix provides both a prediction of how the neuron would respond to that signal, which is accomplished by convolution, and indicates which features of the signal the neuron prefers. Thus, our analysis showed that the majority of the conspecific social calls were composed mostly of ripples moving in the downward spectral direction. Consistent with this feature of the calls is that the STRFs of most IC neurons are directionally selective, with strong preferences for FM signals moving downward at sweep rates that correspond to the FM velocities present in their natural signals. The quantitative agreement between the spectrotemporal properties of bat calls and the spectrotemporal features of ripple stimuli that evoke responses in IC neurons indicates that many IC neurons are tuned to the features of their conspecific vocalizations. These results are consistent with those reported by Woolley et al. (2005) who showed that some spectral and temporal modulations that comprised the STRFs in the midbrain of the zebra finch were also prominent features of the spectral and temporal modulations present in its conspecific song.

Inhibition shapes directional selectivity

Directional selectivity for downward FM sweeps was strongly influenced by inhibition, since blocking inhibition greatly reduced the directional preference for downward compared to upward sweeping ripples. However, it was not inhibition per se that shapes directional selectivity but rather it was the tilting of the inhibitory surrounds along the spectrotemporal axis of the STRF, as shown by the contribution of the linear inseparability of IC neurons to direction selectivity. Inhibition is therefore tuning most

IC neurons for downward direction-selectivity and thereby shapes their responsiveness to specific features present in the species vocalizations.

Neurons in the Bat's IC are Tuned for Conspecific Natural Sounds

Directional selectivity for FM sweeps has been a recurrent theme in studies of the auditory system in a variety of mammals for more than 30 years (Grinnell 1963; Suga 1964; Suga 1968; Clopton and Winfield 1974; Mendelson and Cynader 1985; Phillips, Mendelson et al. 1985; Poon, Chen et al. 1991; Fuzessery 1994; Gordon and O'Neill 1998; Poon and Yu 2000). FM directional selectivity has been of particular interest in studies of bats due to the universal presence of brief downward sweeping FM components in their echolocation calls (Pollak and Casseday 1989; Neuweiler and Schmidt 1993). Directional selectivity for FM signals in bats was first shown by Suga (Suga 1965; Suga and Schlegel 1973), who also was the first to propose that selectivity is shaped by surround inhibition. This theme has been verified in more recent studies (Fuzessery and Hall 1996; Koch and Grothe 1998), which showed that blocking inhibition reduces FM directional properties of neurons in the IC of bats. More recently, neurons selective for FM velocity were reported in the IC of the pallid bat (Fuzessery, Richardson et al. 2006). Direction-selectivity and velocity tuning were also observed in the auditory cortex of bats (Razak and Fuzessery 2006) and rats (Zhang et al., 2003). Both studies showed that directional selectivity, while not initially formed in the cortex, is enhanced at the cortex by both spectral and temporal asymmetries of inhibition.

The results reported here are consistent with, and expand upon the results reported in previous studies of bats and other mammals. Here we did not present electronically

generated FM sweeps that were tailored to the neuron's BF, but rather we used ripple stimuli as well as a suite calls that Mexican free-tailed bats emit for social communication. We showed that most of the communication calls had prominent downward sweeping FM components with velocities reflected by the tilt of their ripple composition. The importance of downward sweeping FM signals for bats is further underscored by the similar spectrotemporal features of the FM sweeps in both their social communication and echolocation calls. Almost all of these signals are brief FM signals that have starting frequencies of 30-40 kHz and sweep downward about an octave in frequency at velocities ranging from 0-250 octaves/s. Thus, the dominance of direction-selective neurons favoring the downward direction with a similar distribution of preferred velocities as in the calls suggests that a substantial portion of the IC population is tuned by inhibition to respond to the features in its conspecific vocalizations.

Chapter 3: Most Informative Features of Natural Calls

As mentioned in the previous chapter, moving ripple stimuli did not evoke a robust neural response from about half of the neurons sampled. Additionally, extracting a meaningful STRF with the ability to accurately predict responses to natural stimuli was only possible in a minority of IC neurons. On the other hand, most IC neurons responded vigorously to conspecific vocalizations with varying degrees of neural selectivity. This has led us to extract the receptive field of these neurons directly from natural communication signals without relying on broadband synthetic stimuli such as moving ripples. However, deriving sensory receptive fields from natural stimuli presented many challenges for scientists in all sensory modalities including vision and audition. These challenges arise from the statistical complexity of natural signals and the nonlinear response properties of sensory neurons. While this has led to the development of novel computational methods that allowed us to derive the relevant features of natural stimuli that sensory neurons are encoding (Theunissen, David et al. 2001; Machens, Wehr et al. 2004; Sharpee, Rust et al. 2004), little is known thus far regarding the actual neural mechanism sensory neurons are using to create their selectivity for particular features of natural stimuli, and specifically stimuli used for social communication.

In this chapter, we extracted the relevant features of natural communication signals that maximized the information between the stimuli presented and their evoked responses in the IC. We further derived the static nonlinearity that maps the stimulus projection onto each feature to the spiking response of the neuron. This allowed us to

gain a better understanding of the feature selectivity exhibited by each IC neuron and to investigate the neural mechanisms that contributed to this selectivity.

Similar to our results described in the previous chapter, the most informative features extracted from IC neurons revealed their tuning for the direction and velocity of spectral motion. Exploring different functional models that could predict the motion selectivity of each IC neuron, we found two distinct neural mechanisms for motion selectivity in the IC. In about half of the neurons sampled, the most informative features were cooperative and exhibited the same properties described in the energy model for the detection of visual motion (Adelson and Bergen 1985). The other half showed tuning for opposing features with opposite spectral directions, indicating a resemblance to the Reichardt correlation-type model described in Chapter 1. This suggests the evolution of common neural mechanisms for motion detection across different sensory modalities.

METHODS

Acoustic Stimuli

Acoustic signals were pure tones, logarithmic frequency-modulated (FM) sweeps as well as species-specific social communication signals. All stimuli had a 0.5 ms rise and fall time constructed using a cosine-squared function.

FM sweeps were centered around the best frequency (BF) of each neuron and swept with different FM velocities either upward or downward on the logarithmic frequency axis. To construct a logarithmic sweep, we defined FM velocity as,

$v = \frac{\log_2(f_1 / f_2)}{\Delta t}$ where f_0 and f_1 are expressed in Hz as the start and end frequencies, respectively, and Δt is sweep duration. Velocity, v, is expressed in octaves/s where a positive value defines an upward moving sweep whereas a negative value indicates a downward sweep. Thus, we can express the instantaneous frequency of the sweep as $f(t) = f_0 2^{vt}$. In order to write the FM sweep as $\sin(\phi(t))$, we have to integrate over time for the instantaneous phase, $\phi(t) = 2\pi \int_0^t f(\tau) d\tau$. Assuming an initial phase of 0°, the logarithmic FM sweep can be described as follows:

$$s(t) = \sin\left(\frac{2\pi f_0 2^{vt}}{v \ln(2)} \right). \qquad (3.1)$$

We used 25 bat social communication calls in this study. The calls were selected from a larger repertoire and were chosen because their acoustic features represent a range of spectrotemporal patterns that are used in a variety of important behavioral contexts (Bohn, Schmidt-French et al. 2008). Each call varied in length from 0.5 to 4 seconds with a sampling rate of 300 kHz. Most calls had their spectrum between 10-80 kHz although some had energy as low as 6 kHz while others had harmonics that went up to 100 kHz. All calls were presented at a mean intensity of 50-70 dB SPL.

Dimensionality Reduction

To derive the relevant features of natural communication signals that drive the response of an auditory neuron, we used dimensionality reduction methods that model the functional relationship between the auditory stimulus and neural response as a cascade of

a set of linear filters and a static nonlinearity (Bialek and de Ruyter van Steveninck 2005). This was done using the Linear-Nonlinear-Poisson (LNP) model (Simoncelli, Pillow et al. 2004), modified to work with natural stimuli as described in (Theunissen, David et al. 2001; Touryan, Felsen et al. 2005), and optimized using information theoretic methods (Pillow and Simoncelli 2006). In this model, the spiking response of a neuron, r, to a given stimulus, s, is modeled as a set of linear filters, $k_1,\ldots,k_n$, with their convolution output run through a static nonlinear function, g, as follows:

$$r(t) = g(k_1 * s, k_2 * s, \ldots, k_q * s) , \qquad (3.2)$$

where q is the number of relevant dimensions that span the feature subspace needed to capture the stimulus-response relationship of the neuron. Here * denotes convolution over time such that $k * s = \int_0^t k(\tau)s(t-\tau)d\tau$, and g is a static nonlinear function that maps q-dimensions onto a spiking rate output, r.

Natural Stimulus Correlations

Each natural sound was converted from a sound pressure waveform to a spectrogram form using a windowed discrete-time Fourier transform with zero mean and log amplitude. The resulting spectrogram for each stimulus segment had n bins in time with a bin-size of around 1 ms and m bins in log frequency with each bin spanning ¼ of an octave. Each stimulus preceding a given time, t, was thus written as a single vector, s_t, with $n \times m$ dimensions. All natural stimuli presented can then be written as $S = [\ s_0^{\mathrm{T}}\ s_1^{\mathrm{T}}\ \ldots\ s_N^{\mathrm{T}}\ ]$, where N is the total number of time samples in the overall spectrogram, and T denotes the vector transpose. Writing each linear filter, k, from

Equation 3.2 in vector form as well enables us to write the convolution operator as a projection across each filter such that $k * S = k^T S$.

In order to use spike-triggered averaging and covariance methods with natural stimuli, the stimulus had to be corrected for its second-order spectrotemporal correlations (Theunissen, David et al. 2001; Touryan, Felsen et al. 2005). First, the stimulus autocorrelation matrix was computed as $A = S^T S$, which was then decomposed into its eigenvectors, U, and eigenvalues $\lambda_1, \ldots, \lambda_n$, using singular value decomposition. Then, the stimuli was "whitened" or "normalized" by correcting for the stimulus correlations as follows:

$$S_w = SU \begin{pmatrix} \frac{1}{\sqrt{\lambda_1}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sqrt{\lambda_c}} \end{pmatrix}, \qquad (3.3)$$

where $c < n$ such that only a subset of the eigenvectors are used for normalization since using very small eigenvalues will result in the amplification of high-frequency noise (Touryan, Felsen et al. 2005; David, Mesgarani et al. 2007; Lesica and Grothe 2008). The percentage of eigenvectors used for whitening the stimulus, often referred to as the cutoff value, was treated as a free parameter and its value was chosen for each neuron such that it maximized the prediction accuracy of the test stimulus. For most neurons the cutoff value was between 30-50%.

Most Informative Subspace

After correcting for stimulus correlations, the spike-triggered average (STA) and covariance (STC) were computed for each neuron. The STA is simply,

$$\mu = \frac{1}{N}\sum_{i=1}^{N} s_i \,, \tag{3.4}$$

where s_i is the stimulus vector preceding the ith spike and N is the total number of spikes evoked by all stimuli. The STC is then derived as

$$C = \frac{1}{N-1}\sum_{i=1}^{N}\left(s_i - \mu\right)\left(s_i - \mu\right)^T \,. \tag{3.5}$$

While the STA and/or the significant eigenvectors of the STC could provide us with the relevant directions in stimulus space that span the feature subspace of the neuron, as described in Equation 3.2, they are restricted to orthogonal directions and it is difficult to know which axes of the subspace are most informative. Instead, we used an information-theoretic approach where both the STA and STC are used to find the most informative subspace that maximizes the information between the raw stimuli and the stimuli that evoked a neural response. This was done using the "information-theoretic spike-triggered average and covariance" (iSTAC) analysis as described by (Pillow and Simoncelli 2006). This method utilizes the Kullback–Leibler (KL) divergence, an information-theoretic measure of the difference between two probability distributions (Cover and Thomas 2006), specifically, the difference between the probability distribution of the raw stimuli, $P(s)$, and the stimuli that evoked a spiking response, $P(s \mid spike)$, such that,

$$D_{KL}[P(s \mid spike) \| P(s)] = \int P(s \mid spike) \log\left(\frac{P(s \mid spike)}{P(s)}\right) ds \qquad (3.6)$$

Assuming both distributions are well approximated by a Gaussian, where the raw stimulus distribution was whitened to have zero mean and unit covariance, as shown above, and the spike-triggered stimulus has a mean and covariance described in the STA and STC, respectively, then the KL divergence within a given subspace, B, can be reduced to the following (Pillow and Simoncelli 2006),

$$D_{KL}(B) = \frac{1}{2}\left(tr\left[B^T\left(C + \mu\mu^T\right)B\right] + \log\left|B^T CB\right| - m\right), \qquad (3.7)$$

where $tr(.)$ and $|.|$ indicate the matrix trace and determinant, respectively. The matrix B that maximizes the above equation gives us the most informative subspace, with its m-columns representing the set of linear filters that span this subspace. Therefore, the KL divergence is optimized as an objective function. First for a 1D subspace, where m=1, then grown incrementally by adding columns to B such that KL divergence is maximized for each dimensionality. At each step, several initialization points are selected from the significant eigenvectors of the STC to ensure the optimization converges to a global maximum. To determine the number of significant subspace dimensions, a nested bootstrap test was used to examine whether the information gained by increasing the dimensionality is significantly above that expected from random sampling.

After finding the most informative subspace for the whitened stimulus, the columns of B, which represent the most informative dimensions ($b_1,\ldots,b_m$), are projected back to the unwhitened space as follows,

$$k_i = b_i^T \begin{pmatrix} \frac{1}{\sqrt{\lambda_1}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sqrt{\lambda_c}} \end{pmatrix} U^{-1} . \qquad (3.8)$$

Static Nonlinearity

After finding the most informative set of linear filters that span the feature subspace of each neuron, as described above, an *m*-dimensional static nonlinearity has to be found that maps the stimulus projection across the most informative subspace onto an actual spiking rate response. When the maximally informative subspace is low dimensional, $m \leq 2$, the nonlinearity can be easily estimated by first projecting the stimulus across the subspace, $s^* = B^T s$, then the nonlinearity can be estimated as follows:

$$P(spike \mid s^*) = \alpha \frac{P(s^* \mid spike)}{P(s^*)}, \qquad (3.9)$$

where α is proportional to the mean spike rate, $P(spike)$. This is often termed the histogram method, where the nonlinearity is derived by taking the ratio of the spike-triggered stimulus to the raw stimulus distributions, both projected across the most informative subspace. For higher dimensions, estimating the full nonlinearity is a lot more involved, but one can still use the histogram method across each informative dimension individually.

Surgical procedures

Experiments were conducted on Mexican free-tailed bats, *Tadarida brasilensis mexicana*, captured from local sources in Austin, Texas. Surgical procedures were as described in a previous report (Xie, Gittelman et al. 2007). In brief, bats were sedated with isoflurane (inhalation) and then anesthetized with an intraperitoneal injection of ketamine/xylazine (75–100 mg/kg ketamine, 11–15 mg/kg xylazine; Henry Schein). Recordings began after recovery from the anesthetic, and thus all data were obtained from awake animals. Water was presented periodically with an eyedropper. Bats typically lay quietly during the experiments. If they showed signs of discomfort, data collection was stopped and doses of the neuroleptic, ketamine hydrochloride (1:40 dilution, 0.01 ml injection) were administered. All experimental procedures were in accordance with a protocol approved by the University of Texas Institutional Animal Care Committee.

Electrophysiology

Single units were recorded with a single micropipette filled with buffered 1 M NaCl and 2% Fast green (pH 7.4) to enhance the visibility of the electrode. The electrode was positioned over the IC and was subsequently advanced from outside of the experimental chamber with a hydraulic micropositioner (2650; Kopf, Tujunga, CA). Recordings were made at depths ranging from ~300 to 1600 μm, which covered most of the dorsoventral extent of the central nucleus of the IC. The electrode was connected via a silver wire to the headstage of a Dagan BVC 700A amplifier with its output digitized by a National Instruments DAQ board (PCI-6259), which was also used for stimulus generation. Data acquisition and stimulus generation were synchronously run using

custom-built software written in Labview (National Instruments, Austin, TX) and Matlab (MathWorks, Natick, MA). Sound was presented in free field from a 3 inch ribbon tweeter (Fountek JP3.0; Madisound Speakers) positioned 40–50° on the side contralateral to the IC from which recordings were made. The speaker was flat ±6 dB from 3 to 80 kHz. Speakers were calibrated with ¼ inch Brüel and Kjær microphones.

RESULTS

This study was based on 136 IC neurons recorded extracellularly in response to natural conspecific communication signals, FM sweeps and tones. The communication calls were selected from a larger repertoire recorded from a local colony of Mexican free-tailed bats while the animals were engaged in a particular behavioral context (Bohn, Schmidt-French et al. 2008). The selected calls were chosen because their acoustic features represented most of the spectrotemporal patterns found in the larger set.

Neural Selectivity to Natural Calls

A spectrogram of two example calls and their recorded responses from 3 IC neurons are shown in Figure 3.1. As can be seen in the figure, IC neurons showed varying degrees of response selectivity to these natural signals. While some neurons responded vigorously to most vocalizations with very little selectivity (Figure 3.1a), others showed a strong response to a particular subset of the calls with little or no response to the others (Figure 3.1b). A third group showed a higher level of selectivity and was only responsive to individual syllables of these calls (Figure 3.1c). In some

neurons, their selectivity to a given call was similar since each neuron would respond to the same syllable of that call, as in the responses to the second syllable of the first call in Figure 3.1b and 1c. These neurons with similar selectivity could thus be tuned for the same stimulus feature. Most neurons in the IC, however, rarely showed a homogeneous response to all vocalizations presented and even though their responses could be similar in one call it is usually different in another (Klug, Bauer et al. 2002; Andoni, Li et al. 2007; Portfors, Roberts et al. 2009; Schneider and Woolley 2010), as in the responses to the second call of Figure 3.1b and 1c. This heterogeneity in selectivity shows that each IC neuron is tuned for different spectrotemporal features of natural calls, and, therefore, the varying levels of neural selectivity in the IC could not be explained by a simple thresholding mechanism; where all neurons are tuned for the same feature and a higher degree of selectivity is produced by having a higher threshold.

Therefore, deriving the relevant features each neuron is encoding by its spiking output could reveal the mechanism involved in creating neural selectivity to natural communication signals in the auditory midbrain.

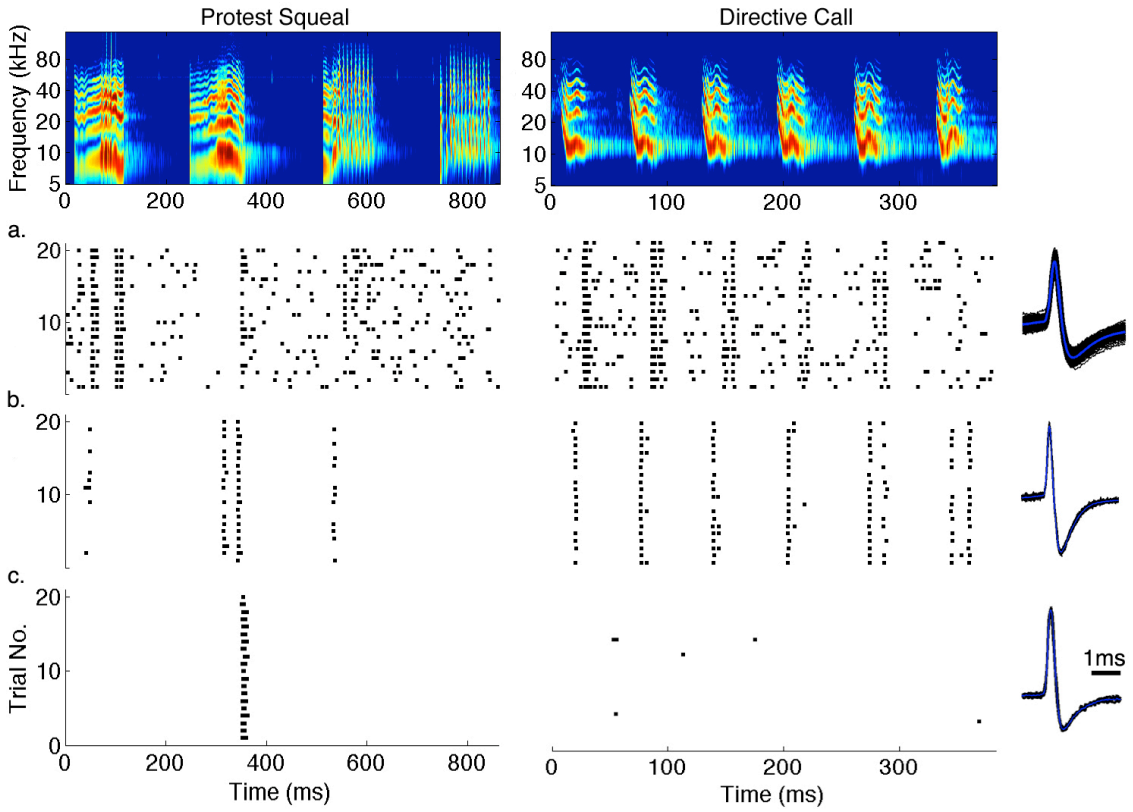

Figure 3.1: Different Levels of Neural Selectivity in the IC.
The top row shows the spectrogram of two example communication calls used by Mexican free-tailed bats. The bottom rows (a-c) display the raster plots of the extracellular response to each call from three IC neurons with their spike waveforms (black) and its average (blue) shown to the right of each row. Each IC neuron displayed a different level of neural selectivity to these calls. Whereas some neurons responded to most syllables in each call (a), others showed a higher degree of selectivity and only a subset of these syllables evoked a neural response (b). Other neurons were even more selective responding to only a single syllable from these two calls (c). This shows that each IC neuron is encoding a different set of spectrotemporal features present in these natural social communication signals.

Most Informative Features and their Static Nonlinearity

In most previous studies, the receptive field of IC neurons was characterized as a single linear filter, which was derived as the average stimulus preceding each spike. It was usually assumed that the spike-triggered average (STA) was linearly related to the response of the neuron and could predict its output by convolving the STA with any arbitrary stimuli. In this manner, the neuron is generally assumed to be tuned for a single feature described by the STA such that the stimulus that is most similar to the STA would predict the largest response and the more the stimulus differs from the STA the weaker is the predicted response. While this was effective in describing the stimulus-response relationship of a minority of neurons in the IC (Andoni, Li et al. 2007{Escabi, 2002 #189; Versnel, Zwiers et al. 2009), the majority of auditory neurons had significant nonlinear response properties and thus the predictions of the STA were relatively poor (Machens, Wehr et al. 2004);Sahani and Linden, 2003).

In this study, we used the linear-nonlinear (LN) cascade model often used in describing the receptive fields of visual neurons (Simoncelli, Pillow et al. 2004; Bialek and de Ruyter van Steveninck 2005; Rust, Schwartz et al. 2005). In this model, it is assumed that the response of a neuron depends on *multiple* spectrotemporal features of the stimulus in addition to the feature described by the STA. These features together could then be treated as a set of linear filters with their outputs run through a multidimensional static nonlinearity. The nonlinearity describes the probability of spike generation as the similarity of the stimulus and each of the features varies (see Methods).

The process of deriving the relevant features and their nonlinearity for an IC neuron is illustrated in Figure 3.2. A subset of the stimuli that preceded each spike is shown in Figure 3.2a, and is referred to as the spike-triggered ensemble (STE). Since we are using natural communication signals that contain strong spectrotemporal correlations which could bias our estimate, the stimuli was first normalized or "whitened" using their second-order correlations as described in Methods. Taking the corrected average of the STE resulted in the STA that is displayed in Figure 3.2b. While the STA could provide us with significant information regarding the feature-selectivity of this neuron, there might be other features this IC cell is tuned for that were not revealed through averaging. To find the complete set of relevant features that resulted in the spiking response of this neuron, one could use two separate methods. The first method involves searching the stimulus space for additional features which when combined together would maximize a quantitative metric such as the predictability of the model (Theunissen, Sen et al. 2000; Machens, Wehr et al. 2004) or the mutual information between the natural stimulus and the spiking response (Sharpee, Rust et al. 2004). An alternative to the latter approach is to use a method similar in some respect to principal components analysis where the spectrotemporal correlations of the STE are computed by deriving the spike-triggered covariance (STC). The set of significant eigenvectors of the STC would then correspond to the relevant linear filters that span the feature subspace of the neuron (Simoncelli, Pillow et al. 2004; Bialek and de Ruyter van Steveninck 2005).

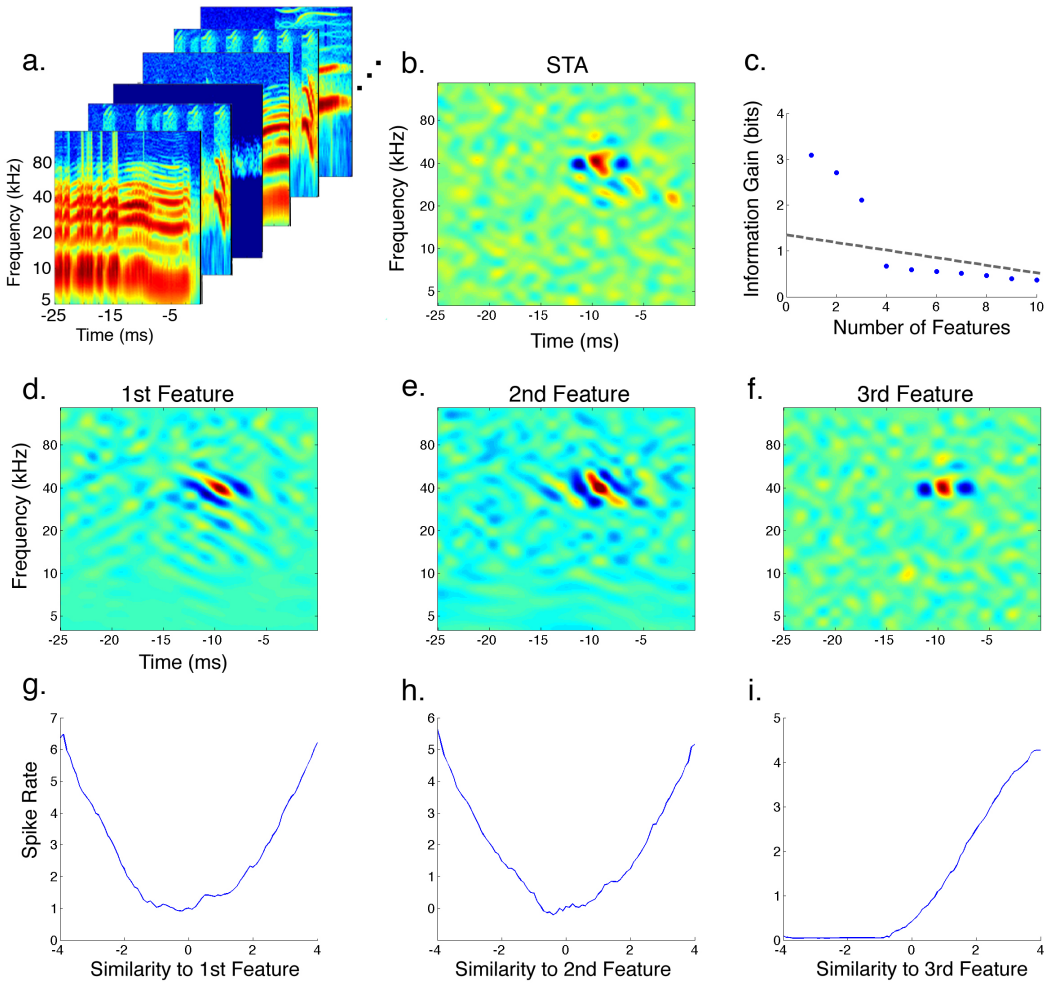

Figure 3.2: Extracting the Most Informative Features and their Static Nonlinearity. To extract the relevant features an IC neuron is encoding in its spiking output, each stimulus segment that preceded a spiking response is collected in the spike-triggered ensemble (STE) shown in (a). Taking the average of the STE after correcting for their spectrotemporal correlations or "whitening" the stimuli resulted in the spike-triggered average (STA) displayed in (b). Using both the STA and the spike-triggered covariance (STC) we searched for the set of spectrotemporal features that maximized the amount of information they preserved between the stimulus and the spiking response. The plot in (c) shows the amount of information gained as the number of features considered is increased. The dashed-line in (c) indicates the level of significant information gain determined by nested bootstrap resampling (see Methods). The three most informative features are shown in the second row (d-f), where the feature ranked as third resembled the STA. The static nonlinearity associated with each feature is displayed in the last row (g-i). Each nonlinearity shows how the spiking probability changes when the similarity between the stimulus and that feature varies.

In this study, we used a hybrid of both methods where we first computed both the STA and STC for each neuron and then used them to search for the set of features that maximized an information-theoretic measure, the Kullbak-Leibler divergence. (Pillow and Simoncelli 2006) (see Methods). This allowed us to derive the most informative stimulus features that each IC neuron is tuned to, where each feature is ranked in the amount of information it preserves about the stimulus-response relationship of the neuron. The three most informative features for the neuron in Figure 3.2 are shown in the second row (Figure 3.2d-f), and Figure 3.2c plots the amount of information gained as the number of features considered is increased. Note that using more than three features for this neuron does not increase the gain in information above the amount expected from noise or undersampling (dashed line). It is also important to note that the STA was not the most informative feature for this neuron but instead was very similar to the feature that was ranked as third. This shows that the STA does not always capture the most significant feature that defines the neural selectivity of an auditory neuron.

The nonlinearity associated with each feature is displayed in the last row (Figure 3.2g-i). Each nonlinearity maps the projection of the stimulus across a particular feature to the spiking rate of the neuron. In other terms, it shows how the spiking rate of the neuron changes depending on the similarity of the stimulus to that feature. To compute the nonlinearity for a given feature, both the raw stimuli as well as the STE were projected across that feature and the ratio of the two distributions resulted in its static nonlinearity. While each plot shows how each feature affects the spiking response of the neuron individually, it does not show the effect of combining the features together. The

full nonlinearity is derived separately and has as many dimensions as the number of relevant features. It defines the spiking probability as the similarity between the stimulus and each of the features changes. In this study, we restricted our analysis to two features since most neurons in our sample were tuned for two significant features, see below, and deriving the full nonlinearity for more than two-dimensions proved to be both computationally involved and sometimes not attainable for the amount of data we had collected. Both symmetric (Figure 3.2g,h), and non-symmetric (Figure 3.2i) nonlinearities were found in the IC. A symmetric nonlinearity indicates that the spiking probability of the neuron increases both when the stimulus is similar to the feature or when it is its complete opposite. Symmetric nonlinearities proved to be important in creating selectivity for spectral motion as discussed below.

Predicting Neural Response

Since the above method allowed us to derive the most informative features that an IC neuron is tuned to, here we investigated how many of these features each cell is encoding and whether using more than one feature could improve our predictability of the spiking response of the neuron. To verify the validity of the derived features, and their associated nonlinearity, they were used to predict the response of each neuron to natural stimuli not used in their derivation. We first used the most informative feature alone and then studied how adding a second feature would improve these predictions. To perform each prediction, the test stimuli were first projected across each feature. A feature projection is generally equivalent to filtering the stimuli with that feature, where a high positive value indicates that the stimulus is very similar in its spectrotemporal shape

to the given feature, and a low negative value indicates that the stimulus has spectrotemporal energy that is in opposite form from that of the feature. Using a one-dimensional (1D) static nonlinearity enables us to map the projection output from one feature into a spiking rate. When two features are used, the two outputs from each feature are translated into a firing rate using a combined 2D nonlinearity. The predicted spiking rate was then compared with the actual firing rate evoked by the test stimuli and a correlation coefficient (CC) was computed.

To evaluate the effect of combining the two most informative features in predicting the neural response in the IC, we first used the first feature alone with its nonlinearity, as shown for an example neuron in the middle row of Figure 3.3. This resulted in a CC value of 0.4 for this neuron. When the second most informative feature was also used with its output combined with the first using the 2D nonlinearity, the CC of the predictions increased to 0.6. The predicted output using the combination of the two most informative features are shown in the bottom row of Figure 3.3. It is evident that this neuron was tuned for multiple spectrotemporal features of the stimulus and using a single feature alone was not enough to produce accurate predictions.

Using the most informative feature alone and its 1D nonlinearity resulted in a mean CC of 0.37 ± 0.24 across the 136 neurons sampled. In a subset of these neurons (49, ~36%), the natural vocalizations presented did not evoke enough spiking responses to derive a meaningful set of features that had significant information gain, and the predictions using the first feature alone were relatively poor (CC < 0.3). Therefore, these neurons were not used for further analysis.

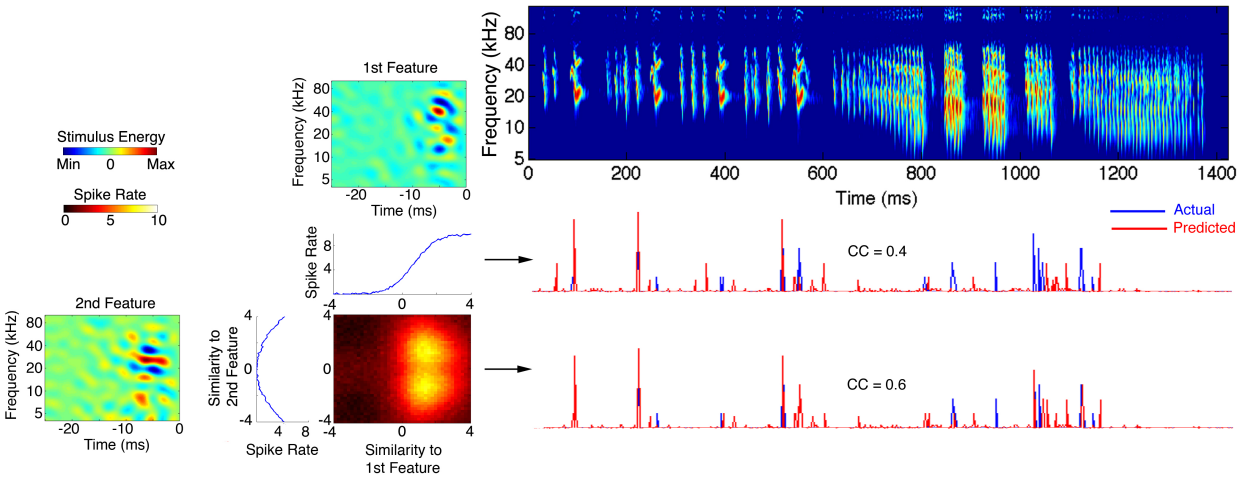

Figure 3.3: Predictions are Improved When Multiple Stimulus Features are Considered.

The two most informative features for an IC neuron are shown on the left, each with its 1D nonlinearity as well as their combined 2D nonlinearity. The 2D nonlinearity shows how the spiking probability varies when the similarity of the stimulus to both features changes. A spectrogram of a bat courtship song not used in deriving the features is displayed on the right with its evoked response (blue) and predicted response (red) shown in the bottom rows. The middle row shows the predicted response using only the most informative feature, and its 1D nonlinearity, resulting in a correlation coefficient (CC) of 0.4. Using both features and their combined 2D nonlinearity resulted in improved predictions with a CC of 0.6. This shows that this IC neuron is tuned for multiple spectrotemporal features of natural signals.

The enhancement of predictability by using a second feature was generally the case for the population of neurons sampled in the IC. The prediction accuracy indicated by the CC value was increased from a mean of 0.46 ± 0.23 to a 0.61 ± 0.15 when both features were used to make predictions. This suggests that neurons in the IC are tuned for multiple features of natural communication signals, which might explain the poor predictions observed in our previous studies that relied solely on the STA.

To evaluate the number of spectrotemporal features each neuron is encoding, we calculated the number of features that produced a significant information gain above that of noise or undersampling. As was shown in Figure 3.2c, the IC neuron had three features that carried significant amount of information that were above noise level and, therefore, these three features and their nonlinearity should fully characterize the receptive field of that neuron and its spectrotemporal tuning. Figure 3.4a shows the number of significant features needed to characterize each neuron in our population. It is evident that the majority of IC cells are tuned for multiple features and only ~7% (6/87) could be fully described by a single spectrotemporal feature that was usually equivalent to the STA.

We further qualitatively evaluated the shape of the static nonlinearity associated with each feature across the neural population sampled. In a minority of neurons, %18 (16/87), the nonlinearity associated with the most informative feature was asymmetric. In these neurons, the most informative feature was equivalent to the STA and the subsequent features had symmetric nonlinearities. The example neuron in Figure 3.3 belonged in that group. However, the majority of these cells, 82% (71/87), had symmetric nonlinearities at least for the two most informative features such as the neuron displayed in Figure 3.2.

Additionally, this group of neurons showed strong selectivity for the direction and velocity of spectral motion as discussed below.

Motion Selectivity

At first glance it could be noted that most of the features encoded by the spiking output of neurons in the bat IC are tilted in shape and are usually spectrotemporally inseparable. Figure 3.4b shows the distribution of inseparability we observed in both the first and second most informative features in our population of IC cells. Since inseparability is generally a prerequisite for direction-selectivity in a linear system, we computed a direction-selectivity index (DSI) for each feature by taking its Fourier transform and comparing the overall power between the two quadrants (Depireux, Simon et al. 2001). A negative DSI indicates selectivity for downward motion whereas a positive value indicates tuning for the upward direction. It was not surprising to see that both features across most neurons were also directionally selective as shown in Figure 3.4c.

To compare the motion selectivity observed in response to natural stimuli to that in response to synthetic stimuli, and to further verify the validity of the extracted features, FM sweeps were presented which varied in both direction and velocity. The FM sweeps were centered around the best frequency (BF) of each neuron and had equal duration but varied in spectral range resulting in different FM velocities in both the upward and downward directions, as illustrated in Figure 3.5a. The DSI computed from responses to FM sweeps across the population of neurons is shown in Figure 3.4d.

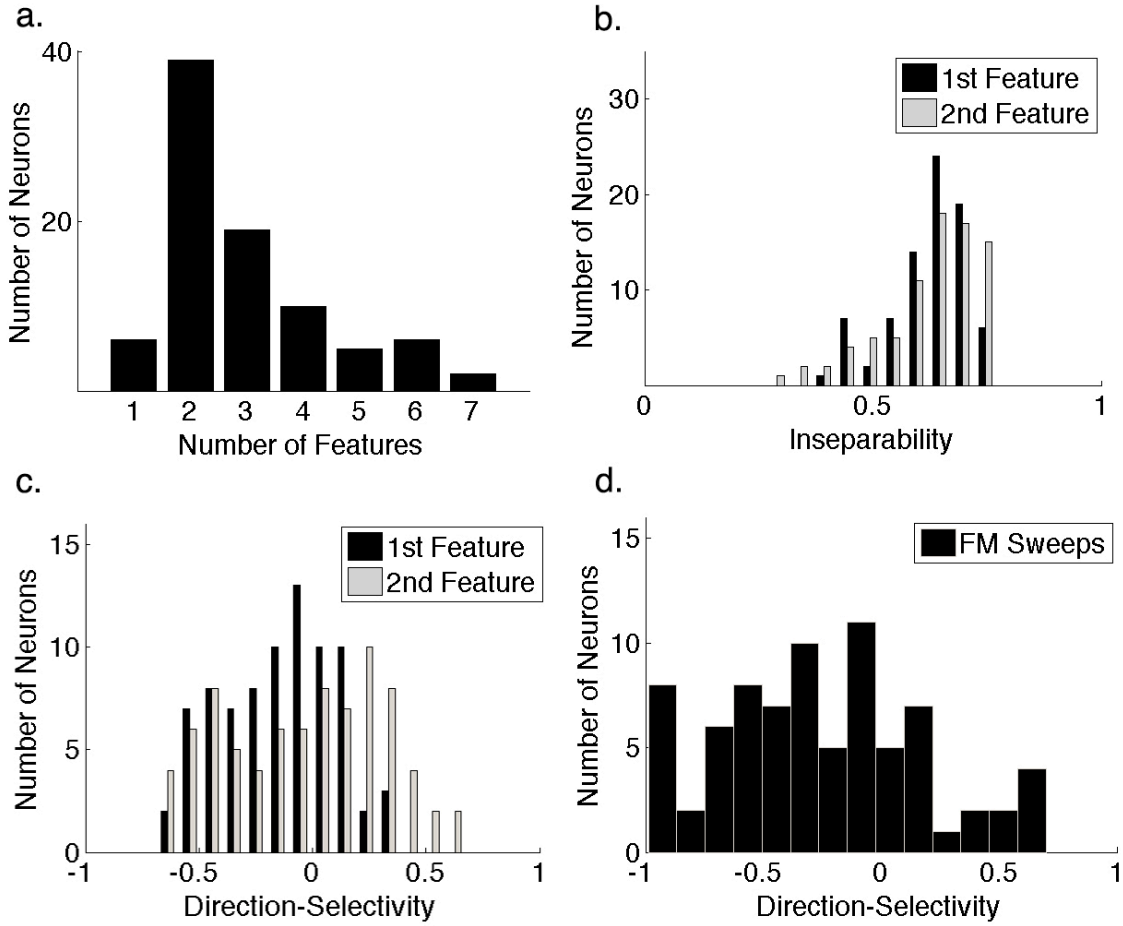

Figure 3.4: Properties of Feature-Selectivity in the IC.
(a) The number of significant features each neuron is encoding shows that the majority of IC neurons in our sample were tuned for two or more spectrotemporal features of the stimulus, and only a minority of them (7%) were tuned for a single feature that was equivalent to the STA. (b) Most features showed strong inseparability since they were usually tilted either upward or downward indicating a preference for the direction of frequency-modulated (FM) sweeps. The mean inseparability index for the first and second most informative features were both around 0.6 (n=87). (c) Direction-selectivity index for the first and second feature shows that they have different directional tuning. While the first feature is biased towards downward (negative) motion with a mean of -0.2, the second feature showed a bimodal distribution with a mean around zero. (d) Comparing direction-selectivity extracted from features of communication signals to selectivity for the direction of synthetic FM sweeps showed similarity to the selectivity observed in the first feature but not the second. Yet selectivity to sweeps had an even stronger bias for the downward direction with a mean of -0.3.

Note that the distribution of DSI in response to sweeps is similar to the DSI computed from the first informative feature, as both distributions show a clear bias for the downward direction. Nevertheless, the DSI calculated from responses to sweeps had an even stronger bias to the downward direction than the first feature, suggesting that the second feature might be playing a role in shaping the motion selectivity of IC neurons. However, the selectivity extracted from the second informative feature was completely different from that of sweeps and showed both downward and upward selectivity across different neurons.

To look closer at motion selectivity for sweeping FM signals in individual neurons, Figure 3.5a shows a raster plot of the responses of an IC neuron to FM sweeps with varying velocities and direction. Note that the neuron only responded to a single FM velocity of -150 octaves/s, where the negative sign denotes the downward direction. The most informative features of this neuron that were extracted from its responses to natural communication signals are displayed in Figure 3.5c,f. Both features were similar to oriented Gabor functions, i.e. a sinusoid with a Gaussian envelope, that were tilted to produce velocity selectivity for the same velocity that evoked the largest response to FM sweeps. The best velocity (BV) of each feature was computed by taking the ratio of the temporal to spectral modulation rates that had a peak magnitude in the Fourier domain (Andoni, Li et al. 2007). The BV for both features was also around -150 octaves/s in agreement with the responses to synthetic FM sweeps. Taking a cross section of each feature perpendicular to its BV shows that each feature is phase shifted from the other by 87°, suggesting that both features are forming a quadrature pair (Figure 3.5e).

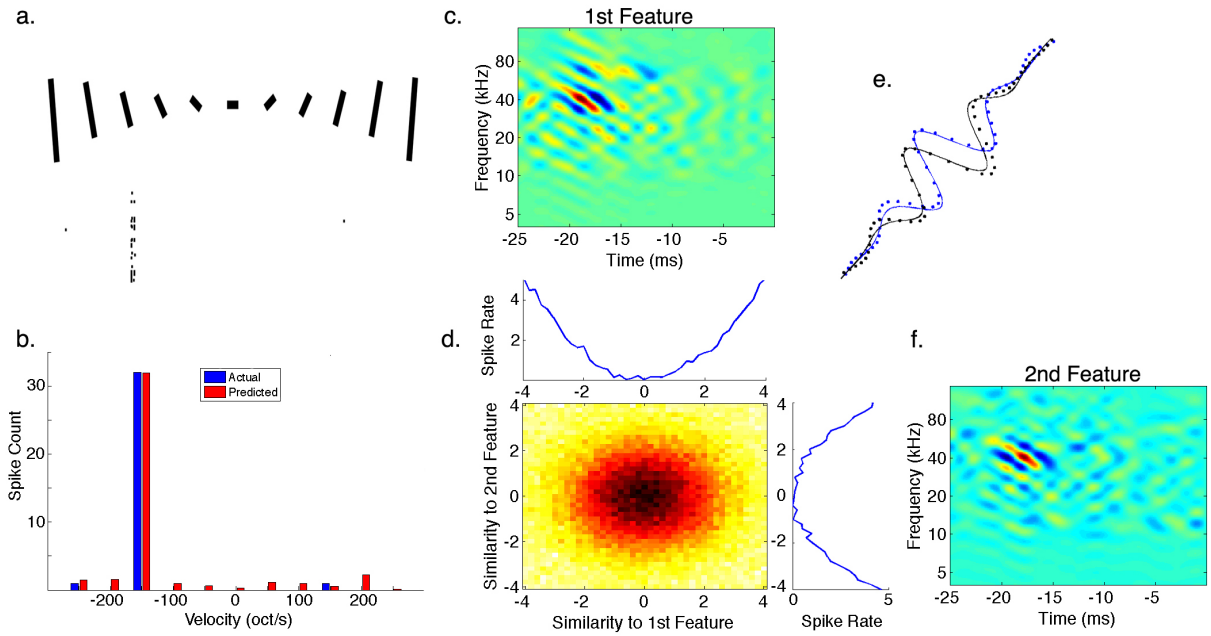

Figure 3.5: Cooperative Features for Motion Selectivity.
(a) A set of downward and upward FM sweeps with varying velocities was presented to an IC neuron with its spiking response displayed below each sweep. (b) Using the most informative features and their nonlinearity extracted from responses to natural stimuli we were able to predict the response of the neuron to only a single sweep velocity of around -150 octaves/second (oct/s) in the downward direction. (c,f) Both features were tilted in the downward direction and had a best velocity (BV) of ~150 oct/s. (d) Both features also had a 1D symmetric nonlinearity and their combined 2D nonlinearity suggests their summation. (e) Fitting a Gabor function to a smoothed cross-section perpendicular to the BV of the first (black) and second (blue) features revealed that they are offset in phase by 87°. This shows that motion selectivity in this IC neuron could be described by a functional model similar to the energy model previously described in the processing of visual motion.

Each feature also had a symmetric 1D nonlinearity with their combined 2D nonlinearity corresponding to their sum (Figure 3.5d). A symmetric nonlinearity increases the spiking probability when the stimulus is either very similar in energy and shape to each feature or that it forms its complete opposite. In this manner the spiking probability is increased only when the stimulus has the corresponding orientation in the spectrotemporal plane. Since both features in this neuron are tuned for the same direction and velocity, their cooperative nonlinear interaction produces strong selectivity for FM sweeping signals. Furthermore, the property of having a quadrature phase-shift together with a symmetric nonlinearity suggests that selectivity for spectral motion in this IC neuron could be explained by a functional model analogous to the energy model previously described in vision (Adelson and Bergen 1985). In this model, the output of two oriented filters, which are phase-shifted to form a quadrature pair, is squared and summed to produce a motion selective output. Figure 3.5b shows the prediction of the model to FM sweep responses and shows that it was able to accurately predict the high degree of selectivity of this neuron to a single FM velocity.

As mentioned previously, not all neurons had their two most informative features tuned for the same direction of motion. In fact around half of the neurons sampled (43/87) had the second feature tuned for the non-preferred direction. An example neuron that is selective for features with opposing directions is shown in Figure 3.6c,f. Although the second feature was tuned to the non-preferred direction, its velocity tuning was very close to the BV of the first feature but in the opposite direction. The BV of both features was 93 octaves/s in opposing directions. When we examine the nonlinearity associated

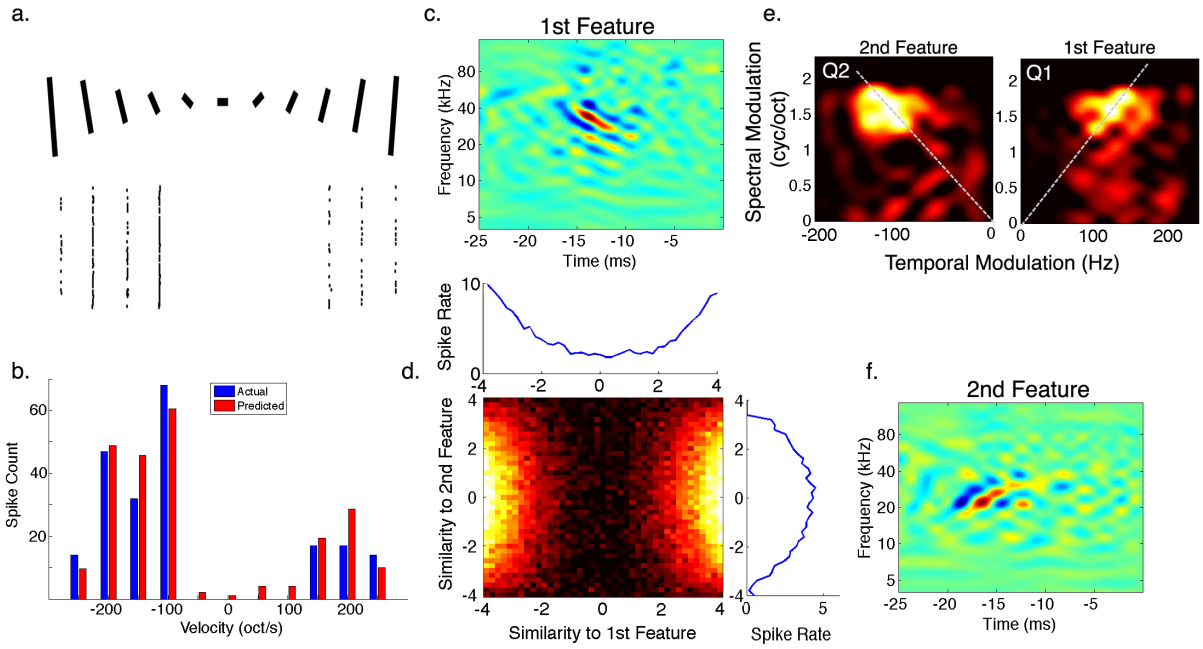

Figure 3.6: Opposing Features for Motion Selectivity.
(a) and (b) are as described in Figure 3.5. (c,f) The most informative features extracted from responses to natural stimuli showed selectivity for opposing FM directions. (d) While both features had a symmetric nonlinearity, the nonlinearity of the second feature was actually suppressive, reducing the response to the upward (non-preferred) direction as shown in the full 2D nonlinearity. (e) Decomposing each feature into its ripple components via a Fourier transform shows that each feature has power within a similar range of spectral and temporal modulations but in opposing quadrants. Furthermore, both features were tuned to the same velocity of 93 oct/s in opposing directions as indicated by the dashed lines.

with each feature (Figure 3.6d), we find that they are both symmetric but the nonlinearity of the second feature is actually suppressive since the spiking probability is decreased when the stimulus is either similar or opposite in shape to that feature. This indicates that the second feature suppressed the response to the non-preferred direction at a velocity close to the BV to which the neuron is tuned in the preferred direction. Figure 3.6e plots the decomposition of both features into their spectral and temporal modulation rates (ripples), via a Fourier transform, which shows that each feature is similar to the mirror image of the other across quadrants, and both showing strong quadrant inseparability a necessary condition for velocity tuning. Using both features and their excitatory and suppressive nonlinearity we were able to predict the response of the neuron to FM sweeps, indicating that our functional model captured the complex velocity tuning of this IC neuron. It is important to note that having an excitatory and suppressive filters tuned in opposite directions are similar to the Reichardt correlation model (Reichardt and Wenking 1969) where two opposing directional subunits produce visual motion selectivity as described in the visual system of the fly (Borst 2000; Bialek and de Ruyter van Steveninck 2005).

Examining the population of IC neurons in our sample showed that half of the cells had both of their most informative features tuned for the same direction and for the same velocity as illustrated by Figure 3.7a. These same features were also phase-shifted by a mean of 93° (Figure 3.7b) indicating a correspondence with the energy model for motion selectivity. The other half of the neurons had features that were tuned to opposing directions with the second feature providing suppression. While their features were tuned

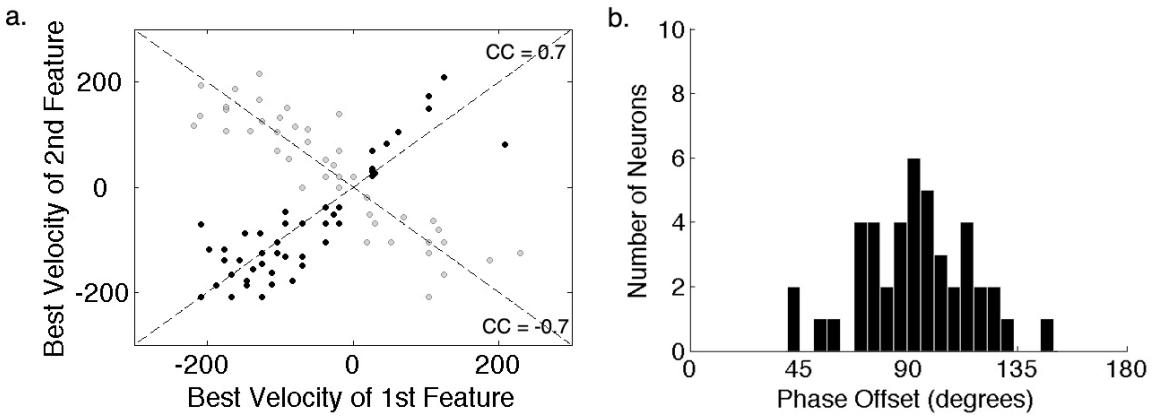

Figure 3.7: Two Distinct Mechanisms for Motion Selectivity in the IC.
(a) Two populations of IC neurons were observed in this study. One with cooperative features tuned for the same direction (black dots, n=44) and the other with opposing features tuned for opposite directions (grey dots, n=43). The best velocity (BV) of the first feature is plotted against the BV of the second and shows that both the cooperative as well as the opposing features were tuned for velocities that are highly correlated. This indicates the importance of spectrotemporal asymmetry in two distinct mechanisms of motion selectivity in the IC. (b) Neurons that were tuned for cooperative features had a phase shift between them with a mean of 93° and a standard deviation of 23°.

for opposite directions, their velocity tuning was very similar (Figure 3.7a). The significance of this similarity in velocity tuning in both excitatory and suppressive features suggests an important role for the spectrotemporal asymmetry in these features and is further discussed in the Discussion section.

Motion in Bat Vocalizations

Our analysis of the spectrotemporal features that IC neurons are encoding has revealed a strong selectivity for the direction and velocity of spectral motion. To understand how this motion selectivity might be playing a role in creating selectivity for the natural communication calls themselves, we analyzed the motion cues present in these signals and compared their modulations over time and frequency to the modulation tuning of IC neurons.

It is apparent from simple visual inspection that most of the communication signals bats emit during different behaviors are mostly composed of frequency modulations that sweep downward or upward at various velocities, see Figure 3.1 and Figure 3.3 for example calls. As described previously in (Andoni, Li et al. 2007), each bat call can be decomposed into its Fourier (ripple) components showing the spectral and temporal modulation rates present in that call (Figure 3.8a,b). This allows us to measure both the FM direction of the call by comparing the power between the two quadrants, and its FM velocity by the alignment of energy around a line with a constant ratio of temporal to spectral modulation rates. Additionally, we can estimate the overall modulation spectrum across all the vocalizations recorded which gives us an overall representation of

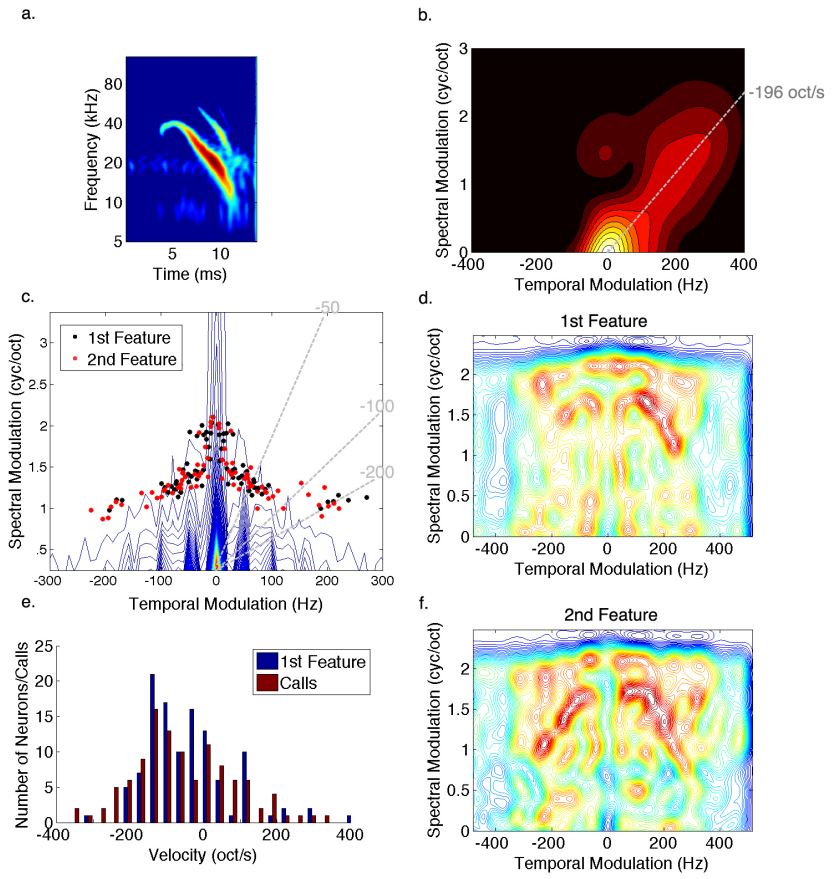

Figure 3.8: Comparing Spectrotemporal Modulations of Conspecific Signals to Neural Tuning.

(a) A spectrogram of a single syllable taken from a courtship call. (b) The Fourier transform of the syllable shows that most of the energy is concentrated around the origin with a tail that is oriented along a line that indicates the sweeping velocity of the syllable. (c) The contour plot shows the modulation spectrum of all bat calls in our repertoire showing 1/*f* distribution that is typical of natural signals. The black and red dots designate the peak modulations present in the first and second most informative features of IC neurons, respectively. Note that peak tuning in the IC is organized to detect various FM velocities (dashed lines) while avoiding redundant energy found in most calls. (d,f) Modulation spectrum of the most informative features shows that IC tuning is specifically aligned around the common modulations in the calls in order to be selective for modulations that represent motion cues found in their conspecific signals such as the extended tail in the modulation spectrum of the above syllable. (e) Distribution of FM velocities found in the calls match the velocities of the most informative feature representing the velocity tuning of IC neurons. This suggests that IC neurons are tuned to detect spectral motion cues present in their conspecific social communication signals.

the modulations in time and frequency that are present across all vocalizations (Singh and Theunissen 2003). We can also apply the same analysis to the features we extracted from neural responses and compare neural tuning in the IC to the acoustic properties of conspecific vocalizations.

Figure 3.8c shows the modulation spectrum of all bat vocalizations. It plots the distribution of spectral and temporal modulations present in the complete repertoire of bat calls. Overlaid on top of the modulation spectrum are the peak modulations found in the most informative features IC neurons are tuned to. One observation that could be made from this plot is that the peak modulation tuning of IC neurons is avoiding the dense areas of the contour plot which indicates that they are avoiding modulations that are most common or redundant across the calls and are tuned instead to modulations that are present in some calls but not others. One can also note that the peak modulation tuning of IC neurons covers a wide range of FM velocities indicated by the dotted lines in the plot. To further examine the modulation tuning of IC neurons, a contour plot of the average modulations found in the first and second most informative features are displayed in Figure 3.8d,f. As was observed from their peak tuning, it is evident that tuning in the IC is aligned to detect modulations that deviate from the common modulations found across calls allowing each neuron to be selective for the modulations that represent a given direction and velocity of spectral motion. In this manner each IC neuron is tuned to respond only for the calls that contain the FM sweeping direction and velocity this neuron is tuned for without responding to calls with modulations outside of its tuning.

Comparing the velocity tuning present in the calls with the velocities IC neurons are tuned for shows a very close agreement (Figure 3.8e). This correspondence between neural tuning and acoustic properties of conspecific communication signals shows that IC neurons are specifically encoding features of these signals through the neural mechanism of motion selectivity.

DISCUSSION

Synthetic vs. Natural Stimuli

Our main motivation for this study was to understand the neural mechanism involved in creating selectivity for specific features of natural communication signals. This required us to derive receptive fields of IC neurons that we were not able to characterize previously with broadband synthetic stimuli such as noise and moving ripples (auditory gratings). In our previous study, we presented a large set of moving ripples with a range of spectral and temporal modulation rates in order to extract a linear estimate of the receptive field of each IC neuron (Andoni, Li et al. 2007). In that study, about half of IC neurons sampled did not respond well to the broadband ripple stimuli, and we were able to extract a meaningful receptive field with accurate predictions in only a small minority of these cells. It is our general observation that neurons in the bat IC do not respond well to broadband stimuli such as noise and ripples, which could be a result of the broad sideband inhibition observed intracellulary (Xie, Gittelman et al. 2007), or simply an outcome of their high degree of selectivity discussed above. Nevertheless, most neurons in the IC respond well to natural stimuli, specifically the conspecific

vocalizations these animals use for social communication. This allowed us to extract the relevant spectrotemporal features each neuron is encoding directly from the communication calls themselves without relying on synthetic stimuli. This strategy comes in agreement with recent studies in birds (Woolley, Gill et al. 2006) and ferrets (David, Mesgarani et al. 2009), where they showed that characterizing the receptive fields of auditory neurons had significant differences depending on whether synthetic or natural stimuli had been used. However, we did present synthetic FM sweeps in this study to verify the validity of the features extracted from natural calls and showed that the selectivity observed with natural stimuli is in agreement to responses to simpler synthetic stimuli.

Tuning for Multiple Spectrotemporal Features of Natural Calls

One of our main findings in this study was that IC neurons are usually tuned for multiple spectrotemporal features of the stimulus. This was evident in our analysis of the number of features required to maximize the amount of information gained between the stimulus and response. It was further verified when the accuracy of the predictions increased when an additional informative feature was used. Since the IC receives a convergence of excitatory and inhibitory projections from various nuclei in the brainstem, it is not surprising to find that most neurons in the IC are actually tuned for multiple spectrotemporal features of sound. This property was studied previously in the processing of binaural cues, which showed that each IC neuron is encoding multiple cues regarding the level and timing differences between the two ears as well as notch detection (Chase and Young 2006). A recent study also showed that tuning for multiple features was also

present in the auditory cortex (Atencio, Sharpee et al. 2008). The result of having multiple features encoded in the spiking output of IC neurons might explain why previous studies had been unsuccessful in producing accurate predictions for a large population of auditory neurons (Machens, Wehr et al. 2004; Andoni, Li et al. 2007); Sahani and Linden, 2003). These studies solely relied on the STA to make these predictions, and as shown above the STA is not always the most informative feature IC neurons are selective for.

Two Distinct Mechanisms for Motion Selectivity in the IC

The majority of IC neurons in our population showed high degree of selectivity to the direction and velocity of acoustic motion across the spectral axis. Our analysis of the most informative features extracted from these neurons revealed two distinct functional mechanisms that enabled these cells to be selective for both the direction and velocity of spectral motion. The first mechanism was analogous to the energy model described in vision (Adelson and Bergen 1985). This model consists of two linear filters (k_1, k_2) that are tilted in the spectrotemporal plane with a quadrature phase-shift. Their output is then squared and summed as follows, $r = \left(k_1^T s\right)^2 + \left(k_2^T s\right)^2$. This enables each neuron to be selective for the FM direction without any spectrotemporal phase-dependence, thus allowing them to be responsive only to stimuli with the correct orientation. By also having filters that were quadrant inseparable, as was shown in Figure 3.6e, they were also selective for the velocity of spectral motion. The energy model for FM selectivity described above agrees with recent intracellular studies in the bat IC where it was shown

that excitation and inhibition are balanced and exhibit similar tuning for the preferred direction (Gittelman, Li et al. 2009).

While the energy model was consistent with about half of the neurons that showed strong selectivity for motion, the other half showed correspondence with a simplified opponent energy model (Adelson and Bergen 1985) as well as the Reichardt correlation model (Reichardt and Wenking 1969). In these models, motion is computed by subunits that are tuned for opposing directions where one increases the neural response whereas the other suppresses it. In this study, this was performed by two linear filters with opposite orientations with motion computed as follows: $r = \left(k_1^T s\right)^2 - \left(k_2^T s\right)^2$. While this could enhance responses to the preferred direction and suppress responses to sweeps moving in the non-preferred direction, it was surprising to see that each filter in these neurons was tuned for the same velocity but in the opposite direction. This might indicate that excitatory and inhibitory inputs innervating these IC neurons have the opposite temporal asymmetry across the frequency axis. In other words, excitatory inputs from different frequency channels might have varying delays to produce coincidence for a particular velocity in the preferred direction, while inhibitory inputs have the opposite delays on the frequency axis thereby suppressing the response for the same velocity but in the non-preferred direction. This scenario would correspond with experimental evidence for the Reichardt model observed in the processing of visual motion in the fly (Borst 2000).

The latter functional model of motion selectivity through opposing filters could be the result of the interaction of excitatory projections from the cochlear nucleus with inhibitory innervations coming from the ventral nucleus of the lateral limniscus and the superior paraolivary nucleus (Pollak, Gittelman et al. 2010). The model is also consistent with recent studies conducted in the auditory cortex of bats where facilitatory excitation was observed for tones with different frequencies and a delay consistent with the best velocity of the neuron (Razak and Fuzessery 2008), and it is also in agreement with intracellular recordings of FM selective neurons in the cortex of rats where excitation and inhibition were shown to have different temporal asymmetries (Ye, Poo et al. 2010).

Neural Tuning and Features of Conspecific Vocalizations

Comparing the general spectrotemporal features of conspecific communication sounds to those IC neurons are selective for revealed that IC cells are tuned to respond to spectral motion cues present in these signals. This was evident in the correspondence between FM velocities found in the calls and those IC neurons are tuned to. This agreed with our previous study where we showed that the receptive fields of IC neurons mapped with moving ripples showed tuning for the FM direction and velocities that match those in the vocalizations (Andoni, Li et al. 2007).

Furthermore, modulation tuning in the IC seemed to be avoiding redundant spectral and temporal modulations that are common among all vocalizations and neurons are instead tuned for modulations that differ from one call to another. This property of IC neurons was previously shown in the midbrain of birds (Woolley, Fremouw et al. 2005), which suggests a common mechanism for the discrimination of natural conspecific

sounds across different species. Looking closer at the modulation tuning of IC neurons in the bat showed that they are aligned across various spectral and temporal modulations allowing them to be tuned for the direction and velocity of spectral motion that distinguish each syllable of a call from another. Therefore, selectivity for spectral motion could be the neural mechanism through which IC neurons of the bat are encoding features of natural communication signals.

Chapter 4: Conclusions and Implications

MOTION SELECTIVITY ACROSS MODALITIES

This research has demonstrated that neural mechanisms for motion selectivity are involved in the encoding of conspecific communication signals in the auditory midbrain. The neural computations involved in this selectivity seem to be in agreement with functional models observed in other sensory modalities such as vision, suggesting the evolution of common neural mechanisms across modalities for performing a similar computation. While both visual and auditory neurons seem to be utilizing similar neural mechanisms for motion detection, their motion selective output indicates opposing behavioral meanings for the animal. Visual neurons that are selective for motion are first found in the primary visual cortex (V1) and this selectivity is further refined in area MT as part of the "where" visual pathway across the parietal lobe (Goodale and Milner 1992; Merigan and Maunsell 1993). The information encoded by the spiking output of these visual neurons specifies the location of a visual object in space and the direction or speed of its spatial motion. In contrast, the visual pathway projecting towards the temporal lobe is involved in the "what" aspect of the stimulus such as object identification (Mishkin, Ungerleider et al. 1983; Maunsell and Newsome 2003). Motion selectivity is only found in the "where" pathway of vision and is not involved in recognizing or identifying a visual object. This is in disparity with our findings that suggest that motion selectivity in audition had evolved to specifically identify an acoustic object as part of auditory scene analysis discussed by Bregman (Bregman 1990)(1990). Since the spiking of each motion selective neuron in the auditory midbrain encodes a certain direction and velocity of

spectral motion, and as we have showed that motion selectivity can accurately predict the selectivity to communication calls, this information could be easily decoded in a downstream nuclei to identify the type of call emitted and then build a perception of its behavioral relevance. Therefore, motion selectivity must be an integral part of the stream segregation performed by the auditory system, and allows the listener to both identify and perceptually group different sounds into coherent representations of objects in the auditory scene (Bregman 1990). Thus, spectral motion in audition must be involved in the "what" auditory pathway in disagreement with its contribution to the "where" pathway in vision.

SPATIAL VS. SPECTRAL MOTION

In addition to spectral motion, an acoustic sound source could also move in space. To compute spatial motion the auditory system must integrate both arrival time and intensity information from both ears. Thus, spatial motion in audition does not lead to physical motion on the sensory epithelium but is rather encoded in the output of both cochleas. Therefore, the neural mechanisms involved in the processing of acoustic spatial motion must be different than mechanisms observed in selectivity for spectral or visual motion.

Jeffres(Jeffress 1948) suggested a functional model for localizing a low frequency sound source based on the timing disparity between the two ears. His model consisted of correlating the inputs from the two ears while having a delay to account for the timing difference that encodes the location of sound in azimuth. In essence his model is comparable to the Reichardt correlation-type model for motion selectivity. However, the

Jeffres model does not compute any motion information but rather encodes the stationary location of a sound source. There have been many studies demonstrating strong experimental evidence for the Jeffres model, for a review see (Joris, Smith et al. 1998), but whether it plays a role in computing spatial motion is still unknown. In fact, the general mechanism for producing selectivity for spatial motion in audition is still unclear, although some studies have shown that neurons in the IC do respond to both the direction and speed of a moving sound source (Wilson and O'Neill 1998; Ingham, Hart et al. 2001; Wagner and von Campenhausen 2002).

NEURAL TUNING FOR FEATURES OF CONSPECIFIC SIGNALS

One of the main findings of this research revealed that neural tuning in sensory systems is specifically tailored for the natural conspecific signals used by the animal, suggesting that sensory neurons have evolved to specifically respond to these signals. While this might not be a surprising result to many, its implication for future research as well as our understanding of sensory systems is imperative. Most early studies of sensory systems have concentrated on simple synthetic stimuli to test the tuning and response properties of sensory neurons without considering the natural stimulus statistics these neurons have evolved to respond to. While this approach has provided us with a strong understanding of many neural mechanisms involved in the processing of sensory stimuli, it does not necessarily explain the neural responses to natural stimuli sensory neurons encounter on a daily basis. Furthermore, recent studies have shown that neural tuning measured with natural stimuli is remarkably different than the tuning observed in response to synthetic signals such as noise (Woolley, Gill et al. 2006; David, Mesgarani

et al. 2009). Since our research demonstrated that sensory neurons are indeed tuned to respond to the natural signals present in the environment of the animal, it is important to understand how responses to natural versus synthetic stimuli are related and how can they help us understand the neural mechanisms involved in encoding the features of the natural signals sensory are tuned for.

SPECTRAL MOTION AMONG OTHER PERCEPTUAL CUES

While this research demonstrated the significance of spectral motion cues in encoding features of natural signals, it does not imply that spectral motion is the only neural and perceptual cue used by the auditory system. FM sounds are abundant in both the echolocation and social communication signals of bats and therefore neurons in their auditory system are tuned for features of these calls, as we have demonstrated. However, FM selectivity is not restricted to the IC of bats, but was found in most mammals including mice (Hage and Ehret 2003), rats (Poon and Yu 2000), cats (Mendelson and Cynader 1985) and monkeys (Versnel, Zwiers et al. 2009). If neural tuning corresponded with the features of conspecific vocalizations emitted by these animals, then the prominence of neurons selective for spectral motion might correlate with the abundance of FM signals in the vocalizations of each animal.

In addition to tuning for spectral motion cues, studies have shown that some neurons in the IC of bats were combination sensitive since they only responded when a specific temporal sequence of signals was presented (Wenstrup, Mittmann et al. 1999). Combination sensitivity was shown to also play a role in encoding species-specific vocalizations (Portfors 2004). Other neurons were shown to be duration tuned, only

responding to stimuli with the exact temporal duration (Casseday, Ehrlich et al. 1994). In addition to these monaural properties, most neurons in the IC show sensitivity to binaural cues such as timing and intensity disparities between the two ears (Chase and Young 2006). Therefore, all these emergent response properties found in the auditory midbrain as a result of the massive convergence of brainstem projections, encode different aspect of the acoustic stimulus which enables the nervous system to build a complete neural representation of the auditory scene.

References

Adelson, E. H. and J. R. Bergen (1985). "Spatiotemporal energy models for the perception of motion." J Opt Soc Am A **2**(2): 284-299.

Andoni, S., N. Li, et al. (2007). "Spectrotemporal receptive fields in the inferior colliculus revealing selectivity for spectral motion in conspecific vocalizations." J Neurosci **27**(18): 4882-4893.

Atencio, C. A., T. O. Sharpee, et al. (2008). "Cooperative nonlinearities in auditory cortical neurons." Neuron **58**(6): 956-966.

Barlow, H. B. and W. R. Levick (1965). "The mechanism of directionally selective units in rabbit's retina." J Physiol **178**(3): 477-504.

Bialek, W. and R. R. de Ruyter van Steveninck (2005). "Features and dimensions: Motion estimation in fly vision." eprint arXiv:q-bio/0505003.

Bohn, K. M., B. Schmidt-French, et al. (2008). "Syllable acoustics, temporal patterns, and call composition vary with behavioral context in Mexican free-tailed bats." J Acoust Soc Am **124**(3): 1838-1848.

Bohn, K. M., B. Schmidt-French, et al. (2009). "Versatility and stereotypy of free-tailed bat songs." PLoS One **4**(8): e6746.

Borst, A. (2000). "Models of motion detection." Nat Neurosci **3 Suppl**: 1168.

Borst, A. (2007). "Correlation versus gradient type motion detectors: the pros and cons." Philosophical Transactions of the Royal Society B: Biological Sciences **362**(1479): 369-374.

Borst, A. and M. Egelhaaf (1989). "Principles of visual motion detection." Trends Neurosci **12**(8): 297-306.

Bregman, A. S. (1990). Auditory scene analysis : the perceptual organization of sound. Cambridge, Mass., MIT Press.

Britt, R. and A. Starr (1976). "Synaptic events and discharge patterns of cochlear nucleus cells. II. Frequency-modulated tones." J Neurophysiol **39**(1): 179-194.

Casseday, J. H., D. Ehrlich, et al. (1994). "Neural tuning for sound duration: role of inhibitory mechanisms in the inferior colliculus." Science **264**(5160): 847-850.

Chase, S. M. and E. D. Young (2006). "Spike-timing codes enhance the representation of multiple simultaneous sound-localization cues in the inferior colliculus." J Neurosci **26**(15): 3889-3898.

Clopton, B. M. and J. A. Winfield (1974). "Unit responses in the inferior colliculus of rat to temporal auditory patterns of tone sweeps and noise bursts." Exp Neurol **42**(3): 532-540.

Cover, T. M. and J. A. Thomas (2006). Elements of information theory. Hoboken, N.J., Wiley-Interscience.

Covey, E., J. A. Kauer, et al. (1996). "Whole-cell patch-clamp recording reveals subthreshold sound-evoked postsynaptic currents in the inferior colliculus of awake bats." J Neurosci **16**(9): 3009-3018.

David, S. V., N. Mesgarani, et al. (2009). "Rapid synaptic depression explains nonlinear modulation of spectro-temporal tuning in primary auditory cortex by natural stimuli." J Neurosci **29**(11): 3374-3386.

David, S. V., N. Mesgarani, et al. (2007). "Estimating sparse spectro-temporal receptive fields with natural stimuli." Network **18**(3): 191-212.

Dawe, L. A., J. R. Platt, et al. (1998). "Spectral-motion aftereffects and the tritone paradox among Canadian subjects." Percept Psychophys **60**(2): 209-220.

Depireux, D. A., J. Z. Simon, et al. (2001). "Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex." J Neurophysiol **85**(3): 1220-1234.

Doupe, A. J. and P. K. Kuhl (1999). "Birdsong and human speech: common themes and mechanisms." Annu Rev Neurosci **22**: 567-631.

Efron, B. and R. Tibshirani (1993). An introduction to the bootstrap. New York, Chapman & Hall.

Egelhaaf, M., A. Borst, et al. (1989). "Computational structure of a biological motion-detection system as revealed by local detector analysis in the fly's nervous system." J Opt Soc Am A **6**(7): 1070-1087.

Escabi, M. A., L. M. Miller, et al. (2003). "Naturalistic auditory contrast improves spectrotemporal coding in the cat inferior colliculus." J Neurosci **23**(37): 11489-11504.

Escabi, M. A. and H. L. Read (2005). "Neural mechanisms for spectral analysis in the auditory midbrain, thalamus, and cortex." Int Rev Neurobiol **70**: 207-252.

Escabi, M. A. and C. E. Schreiner (2002). "Nonlinear spectrotemporal sound analysis by neurons in the auditory midbrain." J Neurosci **22**(10): 4114-4131.

Euler, T., P. B. Detwiler, et al. (2002). "Directionally selective calcium signals in dendrites of starburst amacrine cells." Nature **418**(6900): 845-852.

Ewing, A. W. and H. C. Bennet-Clark (1968). "The Courtship Songs of Drosophila." Behaviour **31**(3/4): 288-301.

Fennema, C. L. and W. B. Thompson (1979). "Velocity determination in scenes containing several moving objects." Computer Graphics and Image Processing **9**(4): 301-315.

Fritz, J., S. Shamma, et al. (2003). "Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex." Nat Neurosci **6**(11): 1216-1223.

Fuzessery, Z. M. (1994). "Response selectivity for multiple dimensions of frequency sweeps in the pallid bat inferior colliculus." J Neurophysiol **72**(3): 1061-1079.

Fuzessery, Z. M. and J. C. Hall (1996). "Role of GABA in shaping frequency tuning and creating FM sweep selectivity in the inferior colliculus." J Neurophysiol **76**(2): 1059-1073.

Fuzessery, Z. M., M. D. Richardson, et al. (2006). "Neural mechanisms underlying selectivity for the rate and direction of frequency-modulated sweeps in the inferior colliculus of the pallid bat." J Neurophysiol **96**(3): 1320-1336.

Gibson, G. and I. Russell (2006). "Flying in tune: sexual recognition in mosquitoes." Curr Biol **16**(13): 1311-1316.

Gittelman, J. X., N. Li, et al. (2009). "Mechanisms underlying directional selectivity for frequency-modulated sweeps in the inferior colliculus revealed by in vivo whole-cell recordings." J Neurosci **29**(41): 13030-13041.

Goodale, M. A. and A. D. Milner (1992). "Separate visual pathways for perception and action." Trends in Neurosciences **15**(1): 20-25.

Gordon, M. and W. E. O'Neill (1998). "Temporal processing across frequency channels by FM selective auditory neurons can account for FM rate selectivity." Hear Res **122**(1-2): 97-108.

Grinnell, A. D. (1963). "The neurophysiology of audition in bats: intensity and frequency parameters." J Physiol **167**: 38-66.

Grzywacz, N. M. and C. Koch (1987). "Functional properties of models for direction selectivity in the retina." Synapse **1**(5): 417-434.

Hage, S. R. and G. n. Ehret (2003). "Mapping responses to frequency sweeps and tones in the inferior colliculus of house mice." European Journal of Neuroscience **18**(8): 2301-2312.

Hildreth, E. C. and C. Koch (1987). "The Analysis of Visual Motion: From Computational Theory to Neuronal Mechanisms." Annual Review of Neuroscience **10**(1): 477-533.

Holmstrom, L., P. D. Roberts, et al. (2007). "Responses to social vocalizations in the inferior colliculus of the mustached bat are influenced by secondary tuning curves." J Neurophysiol **98**(6): 3461-3472.

Hubel, D. H. and T. N. Wiesel (1959). "Receptive fields of single neurones in the cat's striate cortex." J Physiol **148**: 574-591.

Hubel, D. H. and T. N. Wiesel (1962). "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex." J Physiol **160**: 106-154.

Huddleston, W. E. and E. A. DeYoe (2003). "First-order and second-order spectral 'motion' mechanisms in the human auditory system." Perception **32**(9): 1141-1149.

Ingham, N. J., H. C. Hart, et al. (2001). "Spatial receptive fields of inferior colliculus neurons to auditory apparent motion in free field." J Neurophysiol **85**(1): 23-33.

Jeffress, L. (1948). "A place theory of sound localization." Journal of Comparative and Physiological Psychology **41**(1): 35-39.

Joris, P. X., P. H. Smith, et al. (1998). "Coincidence detection in the auditory system: 50 years after Jeffress." Neuron **21**(6): 1235-1238.

Klein, D. J., D. A. Depireux, et al. (2000). "Robust spectrotemporal reverse correlation for the auditory system: optimizing stimulus design." J Comput Neurosci **9**(1): 85-111.

Klug, A., E. E. Bauer, et al. (2002). "Response selectivity for species-specific calls in the inferior colliculus of Mexican free-tailed bats is generated by inhibition." J Neurophysiol **88**(4): 1941-1954.

Koch, U. and B. Grothe (1998). "GABAergic and glycinergic inhibition sharpens tuning for frequency modulations in the inferior colliculus of the big brown bat." J Neurophysiol **80**(1): 71-82.

Kowalski, N., D. A. Depireux, et al. (1996). "Analysis of dynamic spectra in ferret primary auditory cortex. I. Characteristics of single-unit responses to moving ripple spectra." J Neurophysiol **76**(5): 3503-3523.

LeBeau, F. E., M. S. Malmierca, et al. (2001). "Iontophoresis in vivo demonstrates a key role for GABA(A) and glycinergic inhibition in shaping frequency response areas in the inferior colliculus of guinea pig." J Neurosci **21**(18): 7303-7312.

Lesica, N. A. and B. Grothe (2008). "Dynamic spectrotemporal feature selectivity in the auditory midbrain." J Neurosci **28**(21): 5412-5421.

Liberman, A. M. and I. G. Mattingly (1989). "A specialization for speech perception." Science **243**(4890): 489-494.

Limb, J. O. and J. A. Murphy (1975). "Estimating the Velocity of Moving Images in Television Signals." Computer Graphics and Image Processing **4**(4): 311-327.

Lindblom, B. E. and M. Studdert-Kennedy (1967). "On the role of formant transitions in vowel recognition." J Acoust Soc Am **42**(4): 830-843.

Machens, C. K., M. S. Wehr, et al. (2004). "Linearity of cortical receptive fields measured with natural sounds." J Neurosci **24**(5): 1089-1100.

Marmarelis, P. Z. and K. Naka (1972). "White-noise analysis of a neuron chain: an application of the Wiener theory." Science **175**(27): 1276-1278.

Maunsell, J. H. R. and W. T. Newsome (2003). "Visual Processing in Monkey Extrastriate Cortex." Annual Review of Neuroscience **10**(1): 363-401.

Mendelson, J. R. and M. S. Cynader (1985). "Sensitivity of cat primary auditory cortex (AI) neurons to the direction and rate of frequency modulation." Brain Res **327**(1-2): 331-335.

Merigan, W. H. and J. H. R. Maunsell (1993). "How Parallel are the Primate Visual Pathways?" Annual Review of Neuroscience **16**(1): 369-402.

Mindlin, G. B. and R. Laje (2005). The physics of birdsong. Berlin, Springer.

Mishkin, M., L. G. Ungerleider, et al. (1983). "Object vision and spatial vision: two cortical pathways." Trends in Neurosciences **6**: 414-417.

Mittmann, D. H. and J. J. Wenstrup (1995). "Combination-sensitive neurons in the inferior colliculus." Hear Res **90**(1-2): 185-191.

Morris, J. S., C. D. Frith, et al. (1996). "A differential neural response in the human amygdala to fearful and happy facial expressions." Nature **383**(6603): 812-815.

Neuweiler, G. and S. Schmidt (1993). "Audition in echolocating bats." Current Opinion in Neurobiology **3**(4): 563-569.

Palombi, P. S. and D. M. Caspary (1996). "GABA inputs control discharge rate primarily within frequency receptive fields of inferior colliculus neurons." J Neurophysiol **75**(6): 2211-2219.

Perrone, J. A. and A. Thiele (2001). "Speed skills: measuring the visual speed analyzing properties of primate MT neurons." Nat Neurosci **4**(5): 526-532.

Phillips, D. P., J. R. Mendelson, et al. (1985). "Responses of single neurones in cat auditory cortex to time-varying stimuli: frequency-modulated tones of narrow excursion." Exp Brain Res **58**(3): 443-454.

Pillow, J. W. and E. P. Simoncelli (2006). "Dimensionality reduction in neural models: an information-theoretic generalization of spike-triggered average and covariance analysis." J Vis **6**(4): 414-428.

Pollak, G. D. and J. H. Casseday (1989). The neural basis of echolocation in bats. Berlin ; New York, Springer-Verlag.

Pollak, G. D., J. X. Gittelman, et al. (2010). "Inhibitory projections from the ventral nucleus of the lateral lemniscus and superior paraolivary nucleus create directional selectivity of frequency modulations in the inferior colliculus: A comparison of bats with other mammals." Hear Res.

Pollak, G. D. and T. J. Park (1993). "The effects of GABAergic inhibition on monaural response properties of neurons in the mustache bat's inferior colliculus." Hear Res **65**(1-2): 99-117.

Poon, P. W., X. Chen, et al. (1991). "Basic determinants for FM responses in the inferior colliculus of rats." Exp Brain Res **83**(3): 598-606.

Poon, P. W. and P. P. Yu (2000). "Spectro-temporal receptive fields of midbrain auditory neurons in the rat obtained with frequency modulated stimulation." Neurosci Lett **289**(1): 9-12.

Portfors, C. V. (2004). "Combination sensitivity and processing of communication calls in the inferior colliculus of the Moustached Bat Pteronotus parnellii." An Acad Bras Cienc **76**(2): 253-257.

Portfors, C. V., P. D. Roberts, et al. (2009). "Over-representation of species-specific vocalizations in the awake mouse inferior colliculus." Neuroscience **162**(2): 486-500.

Priebe, N. J., C. R. Cassanello, et al. (2003). "The neural representation of speed in macaque area MT/V5." J Neurosci **23**(13): 5650-5661.

Priebe, N. J. and D. Ferster (2005). "Direction selectivity of excitation and inhibition in simple cells of the cat primary visual cortex." Neuron **45**(1): 133-145.

Razak, K. A. and Z. M. Fuzessery (2006). "Neural mechanisms underlying selectivity for the rate and direction of frequency-modulated sweeps in the auditory cortex of the pallid bat." J Neurophysiol **96**(3): 1303-1319.

Razak, K. A. and Z. M. Fuzessery (2008). "Facilitatory mechanisms underlying selectivity for the direction and rate of frequency modulated sweeps in the auditory cortex." J Neurosci **28**(39): 9806-9816.

Reichardt, W. (1961). "Autocorrelation, a principle for evaluation of sensory information by the central nervous system." Principles of Sensory Communications, Ed. W. A. Rosenblith. John Wiley, NY (MIT Press): 303-317.

Reichardt, W. and H. Wenking (1969). "Optical detection and fixation of objects by fixed flying flies." Naturwissenschaften **56**(8): 424-425.

Reid, R. C., R. E. Soodak, et al. (1991). "Directional selectivity and spatiotemporal structure of receptive fields of simple cells in cat striate cortex." J Neurophysiol **66**(2): 505-529.

Remez, R. E., P. E. Rubin, et al. (1981). "Speech perception without traditional speech cues." Science **212**(4497): 947-949.

Riley, J. R., U. Greggers, et al. (2005). "The flight paths of honeybees recruited by the waggle dance." Nature **435**(7039): 205-207.

Rust, N. C., O. Schwartz, et al. (2005). "Spatiotemporal elements of macaque v1 receptive fields." Neuron **46**(6): 945-956.

Schneider, D. M. and S. M. Woolley (2010). "Discrimination of communication vocalizations by single neurons and groups of neurons in the auditory midbrain." J Neurophysiol **103**(6): 3248-3265.

Schuller, G. (1997). "A cheap earphone for small animals with good frequency response in the ultrasonic frequency range." Journal of Neuroscience Methods **71**(2): 187-190.

Sen, K., F. E. Theunissen, et al. (2001). "Feature analysis of natural sounds in the songbird auditory forebrain." J Neurophysiol **86**(3): 1445-1458.

Sharpee, T., N. C. Rust, et al. (2004). "Analyzing neural responses to natural signals: maximally informative dimensions." Neural Comput **16**(2): 223-250.

Shu, Z. J., N. V. Swindale, et al. (1993). "Spectral motion produces an auditory after-effect." Nature **364**(6439): 721-723.

Simmons, J. A., M. B. Fenton, et al. (1979). "Echolocation and pursuit of prey by bats." Science **203**(4375): 16-21.

Simoncelli, E. P., J. Pillow, et al. (2004). Characterization of neural responses with stochastic stimuli. The Cognitive Neurosciences, III. M. Gazzaniga, MIT Press**:** 327-338.

Sinex, D. G. and C. D. Geisler (1981). "Auditory-nerve fiber responses to frequency-modulated tones." Hear Res **4**(2): 127-148.

Singh, N. C. and F. E. Theunissen (2003). "Modulation spectra of natural sounds and ethological theories of auditory processing." J Acoust Soc Am **114**(6 Pt 1): 3394-3411.

Steveninck, R. D. R. V. and W. Bialek (1995). "Reliability and Statistical Efficiency of a Blowfly Movement-Sensitive Neuron." Philosophical Transactions: Biological Sciences **348**(1325): 321-340.

Stevens, K. N. and D. H. Klatt (1974). "Role of formant transitions in the voiced-voiceless distinction for stops." J Acoust Soc Am **55**(3): 653-659.

Suga, N. (1964). "Recovery Cycles and Responses to Frequency Modulated Tone Pulses in Auditory Neurones of Echo-Locating Bats." J Physiol **175**: 50-80.

Suga, N. (1965). "Analysis of frequency-modulated sounds by auditory neurones of echo-locating bats." J Physiol **179**(1): 26-53.

Suga, N. (1968). "Analysis of frequency-modulated and complex sounds by single auditory neurones of bats." J Physiol **198**(1): 51-80.

Suga, N. and P. Schlegel (1972). "Neural attenuation of responses to emitted sounds in echolocating rats." Science **177**(43): 82-84.

Suga, N. and P. Schlegel (1973). "Coding and processing in the auditory systems of FM-signal-producing bats." J Acoust Soc Am **54**(1): 174-190.

Taylor, W. R., S. He, et al. (2000). "Dendritic computation of direction selectivity by retinal ganglion cells." Science **289**(5488): 2347-2350.

Theunissen, F. E., S. V. David, et al. (2001). "Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli." Network **12**(3): 289-316.

Theunissen, F. E., S. V. David, et al. (2001). "Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli." Network: Computation in Neural Systems **12**(3): 289-316.

Theunissen, F. E., K. Sen, et al. (2000). "Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds." J Neurosci **20**(6): 2315-2331.

Theunissen, F. E. and S. S. Shaevitz (2006). "Auditory processing of vocal sounds in birds." Curr Opin Neurobiol **16**(4): 400-407.

Touryan, J., G. Felsen, et al. (2005). "Spatial structure of complex cell receptive fields measured with natural images." Neuron **45**(5): 781-791.

Vater, M. and W. Siefer (1995). "The cochlea of Tadarida brasiliensis: specialized functional organization in a generalized bat." Hear Res **91**(1-2): 178-195.

Versnel, H., M. P. Zwiers, et al. (2009). "Spectrotemporal response properties of inferior colliculus neurons in alert monkey." J Neurosci **29**(31): 9725-9739.

Wagner, H. and M. von Campenhausen (2002). "Distribution of auditory motion-direction sensitive neurons in the barn owl's midbrain." J Comp Physiol A Neuroethol Sens Neural Behav Physiol **188**(9): 705-713.

Waters, C. M. and B. L. Bassler (2005). "Quorum sensing: cell-to-cell communication in bacteria." Annu Rev Cell Dev Biol **21**: 319-346.

Wenstrup, J. J., D. H. Mittmann, et al. (1999). "Inputs to combination-sensitive neurons of the inferior colliculus." J Comp Neurol **409**(4): 509-528.

Wilson, W. W. and W. E. O'Neill (1998). "Auditory motion induces directionally dependent receptive field shifts in inferior colliculus neurons." J Neurophysiol **79**(4): 2040-2062.

Woolley, S. M., T. E. Fremouw, et al. (2005). "Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds." Nat Neurosci **8**(10): 1371-1379.

Woolley, S. M., P. R. Gill, et al. (2006). "Stimulus-dependent auditory tuning results in synchronous population coding of vocalizations in the songbird midbrain." J Neurosci **26**(9): 2499-2512.

Xie, R., J. X. Gittelman, et al. (2007). "Rethinking tuning: in vivo whole-cell recordings of the inferior colliculus in awake bats." J Neurosci **27**(35): 9469-9481.

Xie, R., J. Meitzen, et al. (2005). "Differing roles of inhibition in hierarchical processing of species-specific calls in auditory brainstem nuclei." J Neurophysiol **94**(6): 4019-4037.

Yang, L., G. D. Pollak, et al. (1992). "GABAergic circuits sharpen tuning curves and modify response properties in the mustache bat inferior colliculus." J Neurophysiol **68**(5): 1760-1774.

Ye, C. Q., M. M. Poo, et al. (2010). "Synaptic mechanisms of direction selectivity in primary auditory cortex." J Neurosci **30**(5): 1861-1868.

Zhang, L. I., A. Y. Tan, et al. (2003). "Topography and synaptic shaping of direction selectivity in primary auditory cortex." Nature **424**(6945): 201-205.

Vita

Sari Andoni was born in 1977 in the city of Bethlehem, Palestine. He received his Bachelor of Science degree in Computer Sciences and Mathematics from Brigham Young University in 1999. After joining the laboratory of Dr. George Pollak in 2004 at the University of Texas at Austin, he was able to combine mathematical theory with in vivo electrophysiological experiments in order to study the auditory system of bats. His dissertation work focuses on how the brain processes natural sound, such as animal vocalizations and human speech. In addition to his career in science, he is a talented musician performing locally in Austin, TX as well as various cities around Europe and the Middle East.

Permanent Email Address: andoni@mail.utexas.edu

This dissertation was typed by the author.